W.K. Kellogg Foundation Evaluation Handbook

W.K. Kellogg Foundation Evaluation Handbook

Philosophy and Expectations

Blueprint
Action Steps for Grantees

Planning: Preparing for an Evaluation

1. Identifying Stakeholders and Establishing an Evaluation Team
2. Developing Evaluation Questions
3. Budgeting for an Evaluation
4. Selecting an Evaluator

Implementation: Designing and Conducting an Evaluation

5. Determining Data-Collection Methods
6. Collecting Data
7. Analyzing and Interpreting Data

Utilization: Communicating Findings and Utilizing Results

8. Communicating Findings and Insights
9. Utilizing the Process and Results of Evaluation

Practice

Project evaluations that improve the way projects deliver services, improve project management, and help project directors see problems more clearly and discover new avenues for growth.

January 1998

Dear Reader:

At the W.K. Kellogg Foundation, we believe strongly that evaluation should be conducted not only to demonstrate that a project worked, but also to improve the way it works. An evaluation approach should be rigorous in its efforts to determine the worth of a program and to guide program implementation and management, as well as relevant and useful to program practitioners. Although evaluation is useful to document impact and demonstrate accountability, it should also lead to more effective programs, greater learning opportunities, and better knowledge of what works.

This Evaluation Handbook is designed to encourage dialogue about the role evaluation should play at the project level. We encourage you to think differently about evaluation, so that together we can move the discipline from a stand-alone monitoring process to an integrated and valuable part of program planning and delivery.

We hope that you will find this information to be valuable. At the very least, it should provide a solid base from which to make decisions that ultimately lead to stronger programs and more effective services. In keeping with the spirit of evaluation, we welcome any feedback you care to offer.

Thank you for your interest.

Sincerely,

Anne C. Petersen
Senior Vice President for Programs
W.K. Kellogg Foundation

Foreword by Anne C. Petersen – *Senior Vice President for Programs* . . . . . *I*
Introduction . . . . . *III*

Part One: W.K. Kellogg Foundation's Philosophy and Expectations . . . 1

Chapter 1:
Where We Are: Understanding the W.K. Kellogg Foundation's Framework for Evaluation 2

Chapter 2:
How We Got Here: A Summary of the Evaluation Landscape, History, Paradigms, and Balancing Acts 4
The Evaluation Landscape 4
Historical Context of Evaluation in Human Services 4
The Scientific Method as the Dominant Evaluation Paradigm 5
Balancing the Call to Prove With the Need to Improve 6
Recommendations for a Better Balance 9

Chapter 3:
Three Levels of Evaluation 14
Project-Level Evaluation 14
Cluster Evaluation 17
Program and Policymaking Evaluation 18

Part Two: Blueprint for Conducting Project-Level Evaluation 19

Chapter 4:
Exploring the Three Components of Project-Level Evaluation 20
Context Evaluation 21
Implementation Evaluation 24
Outcome Evaluation 28
Program Logic Model Examples 38

Chapter 5:
Planning and Implementing Project-Level Evaluation 47
Planning Steps: Preparing for an Evaluation 48
Step 1: Identifying Stakeholders and Establishing an Evaluation Team 48
Step 2: Developing Evaluation Questions 51
Step 3: Budgeting for an Evaluation 54
Step 4: Selecting an Evaluator 57
Implementation Steps: Designing and Conducting an Evaluation 69
Step 5: Determining Data-Collection Methods 70
Step 6: Collecting Data 84
Step 7: Analyzing and Interpreting Data 87
Utilization Steps: Communicating Findings and Utilizing Results 96
Step 8: Communicating Findings and Insights 96
Step 9: Utilizing the Process and Results of Evaluation 99

Bibliography *105*
Acknowledgements *110*

Purpose of This Handbook

This handbook is guided by the belief that evaluation should be supportive and responsive to projects, rather than become an end in itself. It provides a framework for thinking about evaluation as a relevant and useful program tool. It is written primarily for project directors who have direct responsibility for the ongoing evaluation of W.K. Kellogg Foundation-funded projects. However, our hope is that project directors will use this handbook as a resource for other project staff who have evaluation responsibilities, for external evaluators, and for board members.

For project staff with evaluation experience, or for those inexperienced in evaluation but with the time and resources to learn more, this handbook provides enough basic information to allow project staff to conduct an evaluation without the assistance of an external evaluator. For those with little or no evaluation experience, and without the time or resources to learn more, this handbook can help project staff to plan and conduct an evaluation with the assistance of an external evaluator.

This handbook is not intended to serve as an exhaustive instructional guide for conducting evaluation. It provides a framework for thinking about evaluation and outlines a blueprint for designing and conducting evaluations, either independently or with the support of an external evaluator/consultant. For more detailed guidance on the technical aspects of evaluation, you may wish to consult the sources recommended in the Bibliography section at the end of the handbook.

Organization of This Handbook

The handbook is made up of two principal sections. Taken together, they serve as a framework for grantees to move from a shared vision for effective evaluation, to a blueprint for designing and conducting evaluation, to actual practice. More specifically:

Part One presents an overview of our philosophy and expectations for evaluation. It includes a summary of the most important characteristics of the Foundation's evaluation approach, to guide all grantees as they plan and conduct project-level evaluation. In addition, Part One reviews the contextual factors that have led to an imbalance in how human service evaluation is defined and conducted, and includes our recommendations for creating a better balance between proving that programs work and improving how they work. Part One ends with an overview of the Foundation's three levels of evaluation, with a particular focus on project-level evaluation (the primary subject of this handbook).

Part Two provides a description of the three components of project-level evaluation that can assist project staff in addressing a broad array of important questions about their project. In addition, Part Two provides our grantees with a blueprint for

planning, designing, and conducting project-level evaluation. This section highlights the important steps to take and links these steps to our philosophy and expectations.

Throughout Part Two, examples are provided in the form of case studies of Foundation grantees. The cases provide project directors with real examples of ways in which evaluation can support projects. The sharing of experiences, insights, what works and doesn't work, and how well the project addresses the needs of people is vital to the learning that takes place within and between Kellogg Foundation-funded projects.

The W.K. Kellogg Foundation was established in 1930 to help people help themselves through the practical application of knowledge and resources to improve their quality of life and that of future generations.

PART ONE CONTENTS

W.K. Kellogg Foundation's Philosophy and Expectations

Chapter 1
Where We Are: Understanding the W.K. Kellogg Foundation's Framework for Evaluation

Chapter 2
How We Got Here: A Summary of the Evaluation Landscape, History, Paradigms, and Balancing Acts
The Evaluation Landscape
Historical Context of Evaluation in Human Services
The Scientific Method as the Dominant Evaluation Paradigm
Balancing the Call to Prove With the Need to Improve
Recommendations for a Better Balance

Chapter 3
Three Levels of Evaluation
Project-Level Evaluation
Cluster Evaluation
Program and Policymaking Evaluation

Evaluation is to help projects become even better than they planned to be.... First and foremost, evaluation should support the project....

W.K. Kellogg Foundation
Evaluation Approach, 1997

W.K. Kellogg Foundation's Philosophy and Expectations

Chapter One

Where We Are: Understanding the W.K. Kellogg Foundation's Framework for Evaluation

The W.K. Kellogg Foundation places a high value on evaluation and has established the following principles to help guide evaluation work.

Strengthen projects: Our goal is to improve the well-being of people. Evaluation furthers this goal by providing ongoing, systematic information that strengthens projects during their life cycle, and, whenever possible, outcome data to assess the extent of change. The evaluation effort should leave an organization stronger and more able to use such an evaluation when outside support ends.

Use multiple approaches: We support multidisciplinary approaches to problem solving. Evaluation methods should include a range of techniques to address important project questions.

Design evaluation to address real issues: We believe community-based organizations should ground their evaluations in the real issues of their respective communities. Therefore, evaluation efforts should also be community based and contextual (based on local circumstances and issues). The primary purpose is to identify problems and opportunities in the project's real communities, and to provide staff and stakeholders with reliable information from which to address problems and build on strengths and opportunities.

Create a participatory process: Just as people participate in project activities, people must participate in project evaluation. The best evaluations value multiple perspectives and involve a representation of people who care about the project. Effective evaluations also prepare organizations to use evaluation as an ongoing function of management and leadership.

Allow for flexibility: We encourage flexibility in the way projects are designed, implemented, and modified. Many Kellogg Foundation-funded projects are not discrete programs, but complex, comprehensive efforts aimed at systemic community change. Therefore, evaluation approaches must not be rigid and

prescriptive, or it will be difficult to document the incremental, complex, and often subtle changes that occur over the life of an initiative. Instead, evaluation plans should take an emergent approach, adapting and adjusting to the needs of an evolving and complex project.

Build capacity: Evaluation should be concerned not only with specific outcomes, but also with the skills, knowledge, and perspectives acquired by the individuals who are involved with the project. We encourage ongoing self-reflection and dialogue on the part of every person involved with evaluation in order to reach increasingly sophisticated understandings of the projects being evaluated. Specifically, the Foundation expects that:

- everyone involved in project evaluation spends time thinking about and discussing how personal assumptions and beliefs affect his or her philosophy of evaluation; and
- everyone (particularly those in leadership positions, such as project directors, evaluators, board members, Kellogg program directors) reflects on the values and politics embedded in the process, and honestly examines how these influence what is focused on and what is missed; who is heard and not heard; how interpretations are made; what conclusions are drawn; and how they are presented.

Our vision for evaluation is rooted in the conviction that project evaluation and project management are inextricably linked. In fact, we believe that "good evaluation" is nothing more than "good thinking."

Effective evaluation is not an "event" that occurs at the end of a project, but is an ongoing process which helps decision makers better understand the project; how it is impacting participants, partner agencies and the community; and how it is being influenced/impacted by both internal and external factors. Thinking of evaluation tools in this way allows you to collect and analyze important data for decision making throughout the life of a project: from assessing community needs prior to designing a project, to making connections between project activities and intended outcomes, to making mid-course changes in program design, to providing evidence to funders that yours is an effort worth supporting.

We also believe that evaluation should not be conducted simply to *prove* that a project worked, but also to *improve* the way it works. Therefore, do not view evaluation only as an accountability measuring stick imposed on projects, but rather as a management and learning tool for projects, for the Foundation, and for practitioners in the field who can benefit from the experiences of other projects.

Chapter Two

How We Got Here: A Summary of the Evaluation Landscape, History, Paradigms, and Balancing Acts

By now you have a good sense of where we stand on evaluation. We take it seriously and follow our philosophy with investments. But where do we stand in the broader landscape? How did we arrive at our particular values and viewpoints about the usefulness and power of evaluation? How did we find our place in the world of evaluation? It takes a short history lesson and some understanding of the art and science of research paradigms to answer these questions.

The Evaluation Landscape

The original mission of program evaluation in the human services and education fields was to assist in improving the quality of social programs. However, for several reasons, program evaluation has come to focus (both implicitly and explicitly) much more on *proving* whether a program or initiative works, rather than on *improving* programs. In our opinion, this has created an imbalance in human service evaluation work—with a heavy emphasis on proving that programs work through the use of quantitative, impact designs, and not enough attention to more naturalistic, qualitative designs aimed at improving programs.

We discuss two reasons for this imbalance:

- the historical context of program evaluation in the U.S.; and
- the influence of the dominant research paradigm on human services evaluation.

Historical Context of Evaluation in Human Services

Although human beings have been attempting to solve social problems using some kind of rationale or evidence (e.g., evaluation) for centuries, program evaluation in the United States began with the ambitious, federally funded social programs of the Great Society initiative during the mid- to late-1960s. Resources poured into these programs, but the complex problems they were attempting to address did not disappear. The public grew more cautious, and there was increasing pressure to provide evidence of the effectiveness of specific initiatives in order to allocate limited resources.

During this period, "systematic evaluation [was] increasingly sought to guide operations, to assure legislators and planners that they [were] proceeding on sound lines and to make services responsive to their public" (Cronbach et al., 1980, pg. 12). One lesson we learned from the significant investments made in the 1960s and '70s was that we didn't have the resources to solve all of our social problems. We needed to target our investments. But to do this effectively, we needed a basis for deciding where and how to invest. "Program evaluation as a distinct field of professional practice was born of two lessons…: First, the realization that there is not enough money to do all the things that need doing; and second, even if there were enough money, it takes more than money to solve complex human and social problems. *As not everything can be done, there must be a basis for deciding which things are worth doing.* Enter evaluation" (Patton, 1997, p. 11).

Today, we are still influenced by this pressure to demonstrate the effectiveness of our social programs in order to ensure funders, government officials, and the public at large that their investments are worthwhile. In fact, since the years of the Great Society, pressure to demonstrate the worth of social programs has increased. Limited resources, increasingly complex and layered social problems, the changing political climate, and a seeming shift in public opinion about the extent to which government and other institutions should support disadvantaged or vulnerable populations have shifted the balance even further to an almost exclusive focus on accountability (prove it works), versus quality (work to improve).

The Scientific Method as the Dominant Evaluation Paradigm

A second factor leading to an emphasis on proving whether a social program works is the influence of the scientific method on human-services evaluation. When most people think about program evaluation, they think of complex experimental designs with treatment and control groups where evaluators measure the impact of programs based on statistically significant changes in certain outcomes; for example, did the program lead to increases in income, improved school performance, or health-status indicators, etc.?

The scientific method is based on hypothetico-deductive methodology. Simply put, this means that researchers/evaluators test hypotheses about the impact of a social initiative using statistical analysis techniques.

Perhaps because this way of conducting research is dominant in many highly esteemed fields and because it is backed by rigorous and well-developed statistical theories, it might dominate in social, educational, and human-services fields—

members of which often find themselves fighting for legitimacy. In addition, this way of doing research and evaluation is well suited to answering the very questions programs/initiatives have historically been most pressured to address: Are they effective? Do they work?

The hypothetico-deductive, natural science model is designed to explain what happened and *show causal relationships* between certain outcomes and the "treatments" or services aimed at producing these outcomes. If designed and conducted effectively, the experimental or quasi-experimental design can provide important information about the particular impacts of the social program being studied. *Did the academic enrichment program lead to improved grades for students? Or increased attendance? Ultimately, was it effective?* However, many of the criteria necessary to conduct these evaluations limit their usefulness to primarily single intervention programs in fairly controlled environments. The natural science research model is therefore ill equipped to help us understand complex, comprehensive, and collaborative community initiatives.

Balancing the Call to Prove With the Need to Improve

Both of these factors—the historical growth in the pressure to demonstrate effectiveness, and the dominance of a research philosophy or model that is best suited to measure change—may have led many evaluators, practitioners, government officials, and the public at large to think of program evaluation as synonymous with demonstrating effectiveness or "proving" the worth of programs. As a result, conventional evaluations have not addressed issues of process, implementation, and improvement nearly as well. And they may very well be negatively impacting the more complex, comprehensive community initiatives (like many of those you operate in your communities) because these initiatives are often ignored as unevaluatable, or evaluated in traditional ways that do not come close to capturing the complex and often messy ways in which these initiatives effect change (Connell, Kubisch, Schorr, Weiss, 1995; Schorr and Kubisch, 1995).

Clearly, demonstrating effectiveness and measuring impact are important and valuable; yet we believe that it is equally important to focus on gathering and analyzing data which will help us improve our social initiatives. In fact, when the balance is shifted too far to a focus on measuring statistically significant changes in quantifiable outcomes, we miss important parts of the picture. This ultimately hinders our ability to understand the richness and complexity of contemporary human-services programs—especially the

system change reform and comprehensive community initiatives which many of you are attempting to implement.

Following are some of the many consequences of operating within a limited evaluation framework:

Consequence 1. We begin to believe that there is only one way to do evaluation. Most people (even those trained in research and evaluation methods) don't realize that methods employed, such as an experimental design, are part of larger world views or paradigms about research. These paradigms are based on different assumptions about:

- What is the nature of reality?
- How do we come to know something?
- What should be the relationship between the researcher/evaluator and the participants in the evaluation process?

The dominant research paradigm described above (hypothetico-deductive), derived from medical and other natural science disciplines, is one such paradigm, but there are others. When one research paradigm begins to dominate a field, it becomes easier to forget that other paradigms—which address different goals and questions—also exist.

Patton explains the effect of forgetting paradigms in this way:

> The very dominance of the hypothetico-deductive paradigm, with its quantitative, experimental emphasis, appears to have cut off the great majority of its practitioners from serious consideration of any alternative evaluation research paradigm or methods. The label "research" [or evaluation] has come to mean the equivalent of employing the "scientific method" of working within the dominant paradigm (1997, pp. 270-271).

In other words, people begin to believe there is only one right way of doing evaluation.

Consequence 2. We do not ask and examine equally important questions. We have already discussed how the dominant research paradigm is suited for addressing certain impact questions—the very questions that, historically, social programs have been pressured to address. However, while it brings certain aspects into focus, it misses other important dimensions of the program.

Here again, research paradigms and philosophies come into play. Even more powerful than the notion that there are different paradigms with different assumptions about the world and how it works (i.e., there is no one right way to do evaluation) is *how much our particular paradigms/assumptions influence the questions we ask; what we think is important to know; the evaluation methods we use; the data we collect; even the interpretations and conclusions we make.*

If we are unaware that evaluation designs and results are based on a paradigm or set of assumptions about how to do evaluation, it is more difficult to see the questions and issues we are missing. These are questions and issues that would come into focus only if we look at the program through the lens of another paradigm.

For example, conventional research methods don't tell us how and why programs work, for whom, and in what circumstances, and don't adequately answer other process and implementation questions. And yet, given the increasingly complex social problems and situations we face today, and the increasingly complex social initiatives and programs developed to solve these problems, these are important questions to address.

Consequence 3. We come up short when attempting to evaluate complex system change and comprehensive community initiatives. This may be the most dangerous consequence of all. In a political and social climate of increasing reluctance to support disadvantaged populations and skepticism about whether any social program works, some of the most promising initiatives are being overlooked and are in danger of being cut off. These are the system change and comprehensive community change initiatives that many know from practice, experience, and even common sense create real change in the lives of children, youth, and families.

However, these initiatives are complex and messy. They do not fit criteria for a "good" quantitative impacts evaluation. There are no simple, uniform goals. There is no standard intervention, or even standard participant/consumer. There is no way to isolate the effects of the intervention because these initiatives focus on integrating multiple interventions.

> And since these initiatives are based on multi-source and multi-perspective community collaborations, their goals and core activities/services are constantly changing and evolving to meet the needs and priorities of a variety of community stakeholders. In short, these initiatives are "unevaluatable" using the dominant natural science paradigm (Connell, Kubisch, Schorr, and Weiss, 1995).

What does this mean? It means that many of these initiatives are not evaluated at all, making it difficult for communities to provide evidence that they are effective. It means that others are evaluated using traditional methods. This leads either to a narrowing of the project to fit the evaluation design (a problem, if what really works is the breadth and multi-pronged nature of these initiatives), or to a traditional impacts report which shows that the initiative had limited impact (because impacts in these complex initiatives may occur over a much longer time period and because many of the critical interim outcomes which are difficult to quantify are overlooked). And it means that a great deal of resources are being wasted and very little is being learned about how these initiatives really work and what their true potential may be (Connell, Kubisch, Schorr, and Weiss, 1995).

Consequence 4. We lose sight of the fact that all evaluation work is political and value laden. When we look at the impacts of a program by using the scientific method only, we miss important contextual factors. This, coupled with the fact that statistical theories can lull us into thinking that we are looking at the neutral and objective truth about the initiative, can mask the fact that evaluation is a political and value-laden process.

Virtually every phase of the evaluation process has political implications which will affect the issues of focus, decisions made, how the outside world perceives the project, and whose interests are advanced and whose are ignored. Evaluators must therefore understand the implications of their actions during all phases of the evaluation and must be sensitive to the concerns of the project director, staff, clientele, and other stakeholders. This understanding requires ongoing dialogue with all groups involved and a responsibility to fully represent the project throughout the evaluation process.

Conflicting agendas, limited funds, different perspectives, or the lack of a common knowledge base may lead to strained relationships between evaluators, project directors, and staff. It is important to talk openly about how these factors affect the evaluation process.

Recommendations for a Better Balance

So, how do we create a better balance, and design evaluations that not only help demonstrate the effectiveness of the project, but also help us know how to improve and strengthen it? The following recommendations form the foundation of our evaluation philosophy:

Recommendation 1. Learn about and reflect on alternative paradigms and methods that are appropriate to our work. As we discussed earlier, conducting research within a single paradigm makes it difficult for us to remember that it is still only one view, and not the only legitimate way to conduct evaluation. There are others—some developed within other disciplines such as anthropology, others developed in reaction to the dominant paradigm. Since we cannot fully describe these complex alternative paradigms here, we provide snapshots of a few to stimulate your thinking.

Interpretivism/Constructivism: The interpretivist or constructivist paradigm has its roots in anthropological traditions. Instead of focusing on explaining, this paradigm focuses on understanding the phenomenon being studied through ongoing and in-depth contact and relationships with those involved (e.g., in-depth observations and interviewing). Relying on qualitative data and rich description which comes from these close, ongoing relationships, the interpretivist/constructivist paradigm's purpose is "the collection of holistic world views, intact belief systems, and complex inner psychic and interpersonal states" (Maxwell and Lincoln, 1990, p. 508). In other words, who are the people involved in the program and what do the experiences mean to them? These holistic accounts are often lost in conventional evaluations, which rely on evaluator-determined categories of data collection, and do not focus on contextual factors.

The primary objective of evaluations based on the assumptions of interpretivism/constructivism is to understand social programs from many different perspectives. This paradigm focuses on answering questions about process and implementation, and what the experiences have meant to those involved. Therefore, it is well suited to helping us understand contextual factors and the complexities of programs—and helping us make decisions about improving project management and delivery.

Feminist Methods: Feminist researchers and practitioners (as well as many ethnic and cultural groups, including African Americans and Hispanics), have long been advocating for changes in research and evaluation based on two principles:

1. Historically, the experiences of girls, women, and minorities have been left out or ignored because these experiences have not fit with developing theories (theories constructed primarily from data on white, middle-class males); and

2. Conventional methodologies, such as the superiority of objective vs. subjective knowing, the distancing of the researcher/evaluator from participants, and the assumptions of value-free, unbiased research/evaluations have been seriously flawed.

Although encompassing a widely diverse set of assumptions and techniques, feminist research methods have been described as "contextual, inclusive, experiential, involved, socially relevant, multi-methodological, complete but not necessarily replicable, open to the environment, and inclusive of emotions and events as experiences" (Nielson, 1990, p. 6, from Reinharz, 1983).

Participatory Evaluation: One research method that is receiving increased utilization in developing countries, and among many of our community-based initiatives, is participatory evaluation, which is primarily concerned with the following: (1) creating a more egalitarian process, where the evaluator's perspective is given no more priority than other stakeholders, including program participants; and (2) making the evaluation process and its results relevant and useful to stakeholders for future actions. Participatory approaches attempt to be practical, useful, and empowering to multiple stakeholders, and help to improve program implementation and outcomes by actively engaging all stakeholders in the evaluation process.

Theory-Based Evaluation: Another approach to evaluation is theory-based evaluation, which has been applied both in the substance abuse area (Chen, 1990) and in the evaluation of comprehensive community initiatives (Weiss, 1995). Theory-based evaluation attempts to address the problems associated with evaluating comprehensive, community-based initiatives and others not well suited to statistical analysis of outcomes. Its underlying premise is that just because we cannot effectively measure an initiative's ultimate outcomes statistically, it does not mean we cannot learn anything about the initiative's effectiveness. In fact, proponents of theory-based evaluation reason that, by combining outcome data with an understanding of the process that led to those outcomes, we can learn a great deal about the program's impact and its most influential factors (Schorr and Kubisch, 1995).

Theory-based evaluation starts with the premise that every social program is based on a theory—some thought process about how and why it will work. This theory can be either explicit or implicit. The key to understanding what really matters about the program is through identifying this theory (Weiss, 1995). This process is also known as developing a program logic model—or picture—describing how the program works. Evaluators and staff can then use this theory of how the initiative effects change to develop key interim outcomes (both for the target population and for the collaborating agencies and organizations) that will lead to ultimate long-term outcomes.

Documenting these interim outcomes (measured in both quantitative and qualitative ways) provides multiple opportunities. It demonstrates whether or not an initiative is on track. Tracking short-term achievements takes some of the

pressure off demonstrating long-term impacts in the first year or two, or having very little to say about the initiative for several years. It allows staff to modify the theory and the initiative based on what they are learning, thereby increasing the potential for achieving long-term impacts. Ultimately, it allows staff to understand and demonstrate effectiveness (to multiple stakeholders) in ways that make sense for these types of complex initiatives.

This evaluation approach also provides a great deal of important information about how to implement similar complex initiatives. What are the pitfalls? What are the core elements? What were the lessons learned along the way?

Recommendation 2. Question the questions. Creating open environments where different perspectives are valued will encourage reflection on which questions are not being addressed and why. Perhaps these questions are hidden by the particular paradigm at work. Perhaps they are not questions that are politically important to those in more powerful positions. Perhaps they hint at potentially painful experiences, not often spoken of or dealt with openly in our society. Encourage staff and the evaluation team to continuously question the questions, and to ask what is still missing. Additionally, review whether you are addressing the following questions:

- How does this program work?
- Why has it worked or not worked? For whom and in what circumstances?
- What was the process of development and implementation?
- What were the stumbling blocks faced along the way?
- What do the experiences mean to the people involved?
- How do these meanings relate to intended outcomes?
- What lessons have we learned about developing and implementing this program?
- How have contextual factors impacted the development, implementation, success, and stumbling blocks of this program?
- What are the hard-to-measure impacts of this program (ones that cannot be easily quantified)? How can we begin to effectively document these impacts?

Recommendation 3. Take action to deal with the effects of paradigms, politics, and values. Perhaps more important than understanding all of the factors that can impact the evaluation process is taking specific actions to deal with these issues,

so that you and your evaluation staff can achieve a fuller understanding of your project and how and why it is working. The following tips can be used by project directors and their evaluation staff to deal with the influence of paradigms, politics, and values:

- Get inside the project—understand its roles, responsibilities, organizational structure, history, and goals; and how politics, values, and paradigms affect the project's implementation and impact.
- Create an environment where all stakeholders are encouraged to discuss their values and philosophies.
- Challenge your assumptions. Constantly look for evidence that you are wrong.
- Ask other stakeholders for their perspectives on particular issues. Listen.
- Remember there may be multiple "right" answers.
- Maintain regular contact and provide feedback to stakeholders, both internal and external to the project.
- Involve others in the process of evaluation and try to work through any resistance.
- Design specific strategies to air differences and grievances.
- Make the evaluation and its findings useful and accessible to project staff and clients. Early feedback and a consultative relationship with stakeholders and project staff leads to a greater willingness by staff to disclose important and sensitive information to evaluators.
- Be sensitive to the feelings and rights of individuals.
- Create an atmosphere of openness to findings, with a commitment to considering change and a willingness to learn.

Each of these areas may be addressed by providing relevant reading materials; making formal or informal presentations; using frequent memos; using committees composed of staff members, customers, or other stakeholders; setting interim goals and celebrating achievements; encouraging flexibility; and sharing alternative viewpoints. These tips will help you deal with political issues, bring multiple sets of values, paradigms and philosophies onto the table for examination and more informed decision making, and will help foster an open environment where it is safe to talk honestly about both the strengths and weaknesses of the project.

Chapter Three

Three Levels of Evaluation

Although the primary focus of this handbook is project-level evaluation, it is important to understand the broad context of Kellogg Foundation evaluation. We have developed three levels of evaluation. Together they maximize our collective understanding and ability to strengthen individual and group projects in grantmaking.

Three Levels of Evaluation

- Project-Level Evaluation
- Cluster Evaluation
- Programming and Policymaking Evaluation

Project-Level Evaluation

Project-level evaluation is the evaluation that project directors are responsible for locally. The project director, with appropriate staff and with input from board members and other relevant stakeholders, determines the critical evaluation questions, decides whether to use an internal evaluator or hire an external consultant, and conducts and guides the project-level evaluation. The Foundation provides assistance as needed. *The primary goal of project-level evaluation is to improve and strengthen Kellogg-funded projects*.

Ultimately, project-level evaluation can be defined as the *consistent, ongoing collection and analysis of information for use in decision making.*

Consistent Collection of Information

If the answers to your questions are to be reliable and believable to your project's stakeholders, the evaluation must collect information in a consistent and thoughtful way. This collection of information can involve individual interviews, written surveys, focus groups, observation, or numerical information such as the number of participants. While the methods used to collect information can and should vary from project to project, the consistent collection of information means having thought through what information you need, and having developed a system for collecting and analyzing this information.

The key to collecting data is to collect it from multiple sources and perspectives, and to use a variety of methods for collecting information. The best evaluations engage an evaluation team to analyze, interpret, and build consensus on the meaning of the data, and to reduce the likelihood of wrong or invalid interpretations.

Use in Decision Making

Since there is no single, "best" approach to evaluation which can be used in all situations, it is important to decide the purpose of the evaluation, the questions you want to answer, and which methods will give you usable information that you can trust. Even if you decide to hire an external consultant to assist with the evaluation, you, your staff, and relevant stakeholders should play an active role in addressing these questions. You know the project best, and ultimately you know what you need. In addition, because you are one of the primary users of evaluation information, and because the quality of your decisions depends on good information, it is better to have "negative" information you can trust than "positive" information in which you have little faith. Again, the purpose of project-level evaluation is not just to prove, but also to improve.

People who manage innovative projects have enough to do without trying to collect information that cannot be used by someone with a stake in the project. By determining who will use the information you collect, what information they are likely to want, and how they are going to use it, you can decide what questions need to be answered through your evaluation.

Project-level evaluation should not be a stand-alone activity, nor should it occur only at the end of a program. Project staff should think about how evaluation can become an integrated part of the project, providing important information about program management and service delivery decisions. Evaluation should be ongoing and occur at every phase of a project's development, from preplanning to start-up to implementation and even to expansion or replication phases. For each of these phases, the most relevant questions to ask and the evaluation activities may differ. What remains the same, however, is that evaluation assists project staff, and community partners make effective decisions to continuously strengthen and improve the initiative.

See Worksheet A on page 16 for highlights of some evaluation activities that might be employed during different phases of project development.

Worksheet A

Possible Project-Level Evaluation Activities

Phase	
Pre-Project	• Assess needs and assets of target population/community. • Specify goals and objectives of planned services/activities. • Describe how planned services/activities will lead to goals. • Identify what community resources will be needed and how they can be obtained. • Determine the match between project plans and community priorities. • Obtain input from stakeholders. • Develop an overall evaluation strategy.
Start-Up	• Determine underlying program assumptions. • Develop a system for obtaining and presenting information to stakeholders. • Assess feasibility of procedures given actual staff and funds. • Assess the data that can be gathered from routine project activities. • Develop a data-collection system, if doing so will answer desired questions. • Collect baseline data on key outcome and implementation areas.
Implementation and Project Modification	• Assess organizational processes or environmental factors which are inhibiting or promoting project success. • Describe project and assess reasons for changes from original implementation plan. • Analyze feedback from staff and participants about successes/failures and use this information to modify the project. • Provide information on short-term outcomes for stakeholders/decision makers. • Use short-term outcome data to improve the project. • Describe how you expect short-term outcomes to affect long-term outcomes. • Continue to collect data on short- and long-term outcomes. • Assess assumptions about how and why program works; modify as needed.
Maintenance and Sustainability	• Share findings with community and with other projects. • Inform alternative funding sources about accomplishments. • Continue to use evaluation to improve the project and to monitor outcomes. • Continue to share information with multiple stakeholders. • Assess long-term impact and implementation lessons, and describe how and why program works.
Replication and Policy	• Assess project fit with other communities. • Determine critical elements of the project which are necessary for success. • Highlight specific contextual factors which inhibited or facilitated project success. • As appropriate, develop strategies for sharing information with policymakers to make relevant policy changes.

Cluster Evaluation

Increasingly, we have targeted our grantmaking by funding groups of projects that address issues of particular importance to the Foundation. The primary purpose for grouping similar projects together in "clusters" is to bring about more policy or systemic change than would be possible in a single project or in a series of unrelated projects. Cluster evaluation is a means of determining how well the collection of projects fulfills the objective of systemic change. Projects identified as part of a cluster are periodically brought together at networking conferences to discuss issues of interest to project directors, cluster evaluators, and the Foundation.

Project directors typically know prior to receiving a grant whether they will be expected to participate in a cluster; but occasionally clusters are formed after grants have been made. Therefore, it is important to be familiar with cluster evaluation even if you are not currently participating in a cluster.

In general, we use the information collected through cluster evaluation to enhance the effectiveness of grantmaking, clarify the strategies of major programming initiatives, and inform public policy debates. *Cluster evaluation is not a substitute for project-level evaluation, nor do cluster evaluators "evaluate" projects.* As stated in the previous section, grantees have responsibility for evaluating their own projects in relationship to their own objectives. Project-level evaluation is focused on project development and outcomes related to the project stakeholders. Cluster evaluation focuses on progress made toward achieving the broad goals of a programming initiative. In short, cluster evaluation looks across a group of projects to identify common threads and themes that, having cross-confirmation, take on greater significance. Cluster evaluators provide feedback on commonalties in program design, as well as innovative methodologies used by projects during the life of the initiative. In addition, cluster evaluators are available to provide technical assistance in evaluation to your project if you request it.

Any data collected by project staff that may be useful to the cluster evaluation should be made available to the cluster evaluator. However, we do not want cluster evaluation to become intrusive to projects nor to drive project-level evaluation. Information is reported to the Foundation in an aggregate form that prevents us from linking data to the individual clients or project participants.

Perhaps the most important aspect of cluster evaluation is that your project will benefit from lessons learned by other similar projects. In turn, what you learn by conducting your project can be of benefit to others.

Program and Policymaking Evaluation

Program and policymaking evaluation is the most macro form of evaluation at the Foundation. Conducted by the Foundation's programming staff, it addresses cross-cutting programming and policy questions, and utilizes information gathered and synthesized from both project-level and cluster evaluation to make effective decisions about program funding and support. This type of evaluation also supports communities in effecting policy change at the local, state, and federal levels.

Taken together, the three evaluation levels provide multiperspective, multisource, multilevel data from which to strengthen and assess individual and groups of projects. The interaction of professionals that occurs across all three levels of evaluation encourages creative and innovative thinking about new ways to evaluate programs and deliver information, which we hope will ultimately lead to sustained positive change at the community level. At the same time, evaluation information from multiple levels, when examined in holistic ways, helps the Kellogg Foundation Board and staff members make effective and informed decisions regarding our programming and policy work.

PART TWO CONTENTS

Blueprint for Conducting Project-Level Evaluation

Chapter 4:
Exploring the Three Components of Project-Level Evaluation
Context Evaluation
Implementation Evaluation
Outcome Evaluation
Program Logic Models

Chapter 5:
Planning and Implementing Project-Level Evaluation
Planning Steps: Preparing for an Evaluation
Step 1: Identifying Stakeholders and Establishing an Evaluation Team
Step 2: Developing Evaluation Questions
Step 3: Budgeting for an Evaluation
Step 4: Selecting an Evaluator
Implementation Steps: Designing and Conducting an Evaluation
Step 5: Determining Data-Collection Methods
Step 6: Collecting Data
Step 7: Analyzing and Interpreting Data
Utilization Steps: Communicating Findings and Utilizing Results
Step 8: Communicating Findings and Insights
Step 9: Utilizing the Process and Results of Evaluation

Knowing is not enough; we must apply.
Willing is not enough; we must do.

— Goethe

Blueprint for Conducting Project-Level Evaluation

Chapter Four

Exploring the Three Components of Project-Level Evaluation: Context, Implementation, and Outcome Evaluation

All too often, conventional approaches to evaluation focus on examining only the outcomes or the impact of a project without examining the environment in which it operates or the processes involved in the project's development. Although we agree that assessing short- and long-term outcomes is important and necessary, such an exclusive focus on impacts leads us to overlook equally important aspects of evaluation—including more sophisticated understandings of how and why programs and services work, for whom they work, and in what circumstances.

By combining the following three components of evaluation developed by leading practitioners and advocated by the Kellogg Foundation—context evaluation, implementation evaluation, and outcome evaluation—project staff will be able to address a broader array of important questions about the project. In our view, a good project evaluation should:

- examine how the project functions within the economic, social, and political environment of its community and project setting (context evaluation);
- help with the planning, setting up, and carrying out of a project, as well as the documentation of the evolution of a project (implementation evaluation); and
- assess the short- and long-term results of the project (outcome evaluation).

Each of the evaluation components focuses on a different aspect of the project. Therefore, evaluation plans should include all three components. How much each component is emphasized, however, depends on the phase of project development, the purpose of the evaluation, and the questions you are attempting to address. Used together, these three components can improve project effectiveness and promote future sustainability and growth.

Context Evaluation: Understanding the Project's Context

Every project is located within a community, and many are part of a larger or umbrella organization. The characteristics of a community and umbrella organization influence a project's plans, how the project functions, and the ability to achieve the project goals. In general, a context evaluation asks: *What about our community and our umbrella organization hinders or helps us achieve project goals? Which contextual factors have the greatest bearing on project successes or stumbling blocks?*

Potential Uses of Context Evaluation

Context evaluation can serve many purposes during the life of a project. Early on, context evaluation might focus on:

- assessing the needs, assets, and resources of a target community in order to plan relevant and effective interventions within the context of the community; and
- identifying the political atmosphere and human services context of the target area to increase the likelihood that chosen interventions will be supported by current community leaders and local organizations.

These types of early evaluation activities often increase community participation, provide motivation for networking among community agencies, and, at times, promote new coalitions.

In later phases of project maturity, context evaluation may focus on:

- gathering contextual information to modify project plans and/or explain past problems (e.g., slower than anticipated growth);
- identifying the political, social, and environmental strengths and weaknesses of both the community and the project; and
- examining the impact of changing federal and state climates on project implementation and success.

Without such information, it will be difficult to make informed decisions about how to improve your project. Furthermore, if environmental barriers to project implementation are understood, seemingly troubled projects might be deemed successful based on the barriers they overcame.

Contextual evaluation is also critical when attempting to replicate programs and services. Oftentimes, even "successful" programs are difficult to replicate because the specific contextual factors (environmental, organizational, human, etc.) that facilitated the program's success were not labeled and understood in the evaluation process.

Focusing a Context Evaluation

For any project, there are multiple contexts which are important to understand. As we have described above, which one(s) to focus on will depend on the phase of the project, the purpose of the evaluation, and the particular evaluation questions you are addressing. Following are two examples of how contextual evaluation can be utilized to improve your project.

Example 1: Community Needs Assessment

During the planning phase of a new program serving welfare mothers, a context evaluation might focus on the demographics of welfare mothers in the community and their access to, and opportunities for, services. In addition, context evaluation questions would focus on the social, economic, and political situation in the community, and of the welfare women as a subgroup within the community. Using this type of information during project planning phases helps to ensure that relevant and culturally sensitive program activities or services are implemented.

A simple and effective way to begin the process of context evaluation during this planning phase is through mapping community needs and assets. This process can help you:

- identify existing community action groups and understand the history of their efforts;
- identify existing formal, informal, and potential leaders;
- identify community needs and gaps in services;
- identify community strengths and opportunities;
- understand your target population (both needs and assets) in order to improve, build, and secure project credibility within the community; and
- create a momentum for project activities by getting community input.

Mapping community needs and assets can also help determine the appropriateness of project goals and provide baseline data for later outcome evaluations. A formal needs assessment can be both time-consuming and

resource-intensive; most projects, however, have the capability to perform an informal or simplified needs assessment.

Example 2: Organizational Assessment

A program that has been up and running for a year and facing difficulties in continuing to serve its target population might focus on organizational contextual factors particular to the program (e.g., leadership styles; staff characteristics such as training, experience, and cultural competence; organizational culture; mission; partner agencies). Such contextual information, when collected and analyzed carefully, can help staff identify stumbling blocks and obstacles to program success and improvement. It helps us better understand why something worked or didn't work, why some outcomes were achieved and others were not.

Through an organizational assessment, project staff can examine the internal dynamics of a project to see how these dynamics may be hindering or supporting project success. Questions to be addressed might include:

- What are the values or environment of the project (internal) and its larger institutional context (umbrella organization)? How are they the same? How do differences in values impede project activities?
- What are the fiduciary, physical space, and other collaborative and administrative relationships between the project and its umbrella institution? How do they relate to project accomplishments or failures? For a proposed activity, are these arrangements adequate?
- What is the structure and size of the project in relation to that of the umbrella organization?
- How does the leadership and organizational structure of the project influence its effectiveness? What is the complexity of the organizational chart? Do organizational decision-making bodies impede or strengthen ongoing or proposed activities?
- What are the characteristics of project staff and leadership? How are project members recruited? What is the organizational culture?
- What resources (e.g., funding, staffing, organizational and/or institutional support, expertise, and educational opportunities) are available to the project and to the evaluation?
- To what extent are opportunities to participate in the evaluation process available for people who have a stake in the project's outcome?

If an organizational assessment does not help to fully explain the project's strengths and weaknesses in serving its target population, another contextual area to examine might be changing federal and state climates and how these climates may be impacting the community and project.

A Final Note: Examining the external and internal contextual environments of a project provides the groundwork for implementation and outcome evaluation. It helps to explain why a project has been implemented the way it has, and why certain outcomes have been achieved and others have not. Evaluating the multiple contexts of a project may also point to situations that limit a project's ability to achieve anticipated outcomes, or lead to the realization that specific interventions and their intended outcomes may be difficult to measure or to attribute to the project itself.

Implementation Evaluation: Understanding How the Project Was Implemented

Implementation evaluation activities enhance the likelihood of success by providing indications of what happened and why. Successful implementation of new project activities typically involves a process of adapting the ideal plan to local conditions, organizational dynamics, and programmatic uncertainties. This process is often bumpy, and in the end, actual programs and services often look different from original plans. Even well-planned projects need to be fine-tuned in the first months of operation, and often information needs to be continually analyzed to make improvements along the way.

Every project director has used implementation evaluation, whether or not they have labeled it as such. Implementation evaluations focus on examining the core activities undertaken to achieve project goals and intended outcomes. Questions asked as part of an implementation evaluation include: *What are the critical components/activities of this project (both explicit and implicit)? How do these components connect to the goals and intended outcomes for this project? What aspects of the implementation process are facilitating success or acting as stumbling blocks for the project?*

Potential Uses of Implementation Evaluation

Implementation evaluation addresses a broad array of project elements. Some potential purposes include:

- identifying and maximizing strengths in development;
- identifying and minimizing barriers to implementing activities;
- determining if project goals match target population needs;
- assessing whether available resources can sustain project activities;
- measuring the performance and perceptions of the staff;
- measuring the community's perceptions of the project;
- determining the nature of interactions between staff and clients;
- ascertaining the quality of services provided by the project;
- documenting systemic change; and
- monitoring clients' and other stakeholders' experiences with the project, and their satisfaction with and utilization of project services.

Focusing an Implementation Evaluation

As with context evaluation, the focus of an implementation evaluation will vary depending on the phase of the project, the purpose of the evaluation, and the particular questions you are attempting to address. Following are three examples of implementation evaluation.

Example 1: New Programs

An implementation evaluation designed for a new or rapidly changing organization might focus on information that would assist decision makers in documenting the project's evolution, and continually assessing whether modifications and changes are connected to goals, relevant contextual factors, and the needs of the target population. To help a project do this, the evaluator must understand, from multiple perspectives, what is happening with the project. How is it being implemented, and why have particular decisions been made along the way? In short, to what extent does the project look and act like the one originally planned? Are the differences between planned and actual implementation based on what made sense for the clients and goals of the project? How is the project working now and what additional changes may be necessary?

Specific questions might include:

- What characteristics of the project implementation process have facilitated or hindered project goals? (Include all relevant stakeholders in this

discussion, such as clients/participants, residents/consumers, staff, administrators, board members, other agencies, and policymakers.)

- Which initial strategies or activities of the project are being implemented? Which are not? Why or why not?
- How can those strategies or activities not successfully implemented be modified or adapted to the realities of the project?
- Is the project reaching its intended audience? Why or why not? What changes must be made to reach intended audiences more effectively?
- What lessons have been learned about the initial planned program design? How should these lessons be utilized in continually revising the original project plan? Do the changes in program design reflect these lessons or other unrelated factors (e.g., personalities, organizational dynamics, etc.)? How can we better connect program design changes to documented implementation lessons?

Example 2: Established Programs

For a program that has been up and running for several years, the implementation evaluation might be designed as a continuous monitoring, feedback, and improvement loop. This type of continual monitoring provides project staff with ongoing feedback to help them recognize which activities are working and which activities need modification or restructuring. Examples of questions addressed in this implementation evaluation include:

- Which project operations work? Which aren't working? Why or why not?
- What project settings (facilities, scheduling of events, location, group size, transportation arrangements, etc.) appear to be most appropriate and useful for meeting the needs of clients?
- What strategies have been successful in encouraging client participation and involvement? Which have been unsuccessful?
- How do the different project components interact and fit together to form a coherent whole? Which project components are the most important to project success?
- How effective is the organizational structure in supporting project implementation? What changes need to be made?

Example 3: Piloting Future Programs

For a project director who is thinking about future growth opportunities, an implementation evaluation might be designed to pilot new ideas and determine if

these ideas make sense and are achievable. This evaluation design might include up-front data gathering from clients in order to delineate more effectively future goals and plans based on still unmet needs or gaps in the services currently provided. Specific questions for this implementation evaluation might include:

- What is unique about this project?
- What project strengths can we build upon to meet unmet needs?
- Where are the gaps in services/program activities? How can the project be modified or expanded to meet still unmet needs?
- Can the project be effectively replicated? What are the critical implementation elements? How might contextual factors impact replication?

Summary

A project implementation evaluation should include the following objectives: Improve the effectiveness of current activities by helping initiate or modify initial activities; provide support for maintaining the project over the long term; provide insight into why certain goals are or are not being accomplished; and help project leaders make decisions. In addition, implementation evaluations provide documentation for funders about the progress of a project, and can be used for developing solutions to encountered problems.

Evaluating project implementation is a vital source of information for interpreting results and increasing the power and relevance of an outcome evaluation. Knowing *why* a project achieves its goals is more important than just knowing that it does. An outcome evaluation can tell you what impact your program/service had on participants, organizations, or the community. An implementation evaluation allows you to put this outcome data in the context of what was actually done when carrying out the project. In fact, without knowing exactly what was implemented and why, it is virtually impossible to select valid effectiveness measures or show causal linkages between project activities and outcomes.

Outcome Evaluation: Determining Project Outcomes

Outcome evaluation is another important feature of any comprehensive evaluation plan. It assesses the short- and long-term results of a project and seeks to measure the changes brought about by the project. Outcome evaluation questions ask: *What are the critical outcomes you are trying to achieve? What impact is the project having on its clients, its staff, its umbrella organization, and its community? What unexpected impact has the project had?*

Because projects often produce outcomes that were not listed as goals in the original proposal, and because efforts at prevention, particularly in complex, comprehensive, community-based initiatives, can be especially difficult to measure, it is important to remain flexible when conducting an outcome evaluation. Quality evaluations examine outcomes at multiple levels of the project. These evaluations focus not only on the ultimate outcomes expected, but also attempt to discover unanticipated or important interim outcomes.

Potential Uses of Outcome Evaluation

Outcome evaluation can serve an important role during each phase of a project's development. Early on, you might focus outcome evaluation on:

- determining what outcomes you expect or hope for from the project; and
- thinking through how individual participant/client outcomes connect to specific program or system-level outcomes.

These types of early evaluation activities increase the likelihood that implementation activities are linked to the outcomes you are trying to achieve, and help staff and stakeholders stay focused on what changes you are really attempting to make in participants' lives.

In later phases of project maturity, an effective outcome evaluation process is critical to:

- demonstrating the effectiveness of your project and making a case for its continued funding or for expansion/replication;
- helping to answer questions about what works, for whom, and in what circumstances, and how to improve program delivery and services; and

- determining which implementation activities and contextual factors are supporting or hindering outcomes and overall program effectiveness.

In the following sections, we provide a range of information about outcome evaluation, along with some of the latest thinking about evaluating project outcomes, particularly for more complex, comprehensive, community-wide initiatives.

Types of Outcomes

Each project is unique and is aimed at achieving a range of different outcomes. The following provides a framework for thinking about the different levels of outcome when developing your outcome evaluation plan.

Individual, Client-Focused Outcomes:

When people think about outcomes, they usually think about program goals. The problem is that, often, program goals are stated in terms of service delivery or system goals (e.g., reduce the number of women on welfare), rather than on clear outcome statements about how clients' lives will improve as a result of the program. Yet when we think about the purposes of social and human services programs, we realize that the most important set of outcomes are individual client/participant outcomes. By this, we mean, "What difference will this program/initiative make in the lives of those served?" When you sit down with program staff to answer this question, it will become clear that "reducing the number of women on welfare" is not a client-focused outcome; it is a program- or system-focused outcome.

There are multiple ways to reduce the number of women on welfare (the stated outcome), but not all are equally beneficial to clients. The program might focus on quick-fix job placement for women into low-skill, low-paying jobs. However, if what many clients need is a long-term skill-building and support program, this method of "reducing the number of women on welfare" might not be the most appropriate or most beneficial program for the clients served.

If we change the outcome statement to be client-focused, we see how it helps us focus on and measure what is truly important to improving the lives of women on welfare. For example, the primary individual-level outcome for this program might be: "Clients will gain life and job skills adequate to succeed in their chosen field," or "Clients will gain life and job skills necessary to be self-reliant and economically independent."

The type of outcomes you may be attempting to achieve at the individual client level might include changes in circumstances, status, quality of life or functioning,

attitude or behavior, knowledge, and skills. Some programs may focus on maintenance or prevention as individual client outcomes.

Program and System-Level Outcomes:
Our emphasis on client-focused outcomes does not mean that we do not care about program and system-level outcomes. You do need to think through what outcomes you are trying to achieve for the program and for the broader system (e.g., improved access to case management, expanded job placement alternatives, strengthened interagency partnerships); however, these outcomes should be seen as strategies for achieving ultimate client/participant outcomes. Once you have determined individual client outcomes, then you can determine which specific program and system-level outcomes will most effectively lead to your stated client improvements. Program and system-level outcomes should connect to individual client outcomes, and staff at all levels of the organization should understand how they connect, so they do not lose sight of client-level outcomes and focus on program outcomes, which are easier to measure and control.

> ***Example:*** An initiative aimed at improving health-care systems by strengthening local control and decision making, and restructuring how services are financed and delivered, has as its core individual, client-centered outcome: "improved health status for those living in the community served." However, it quickly became clear to staff and key stakeholders that the road to improved health status entailed critical changes in health-care systems, processes, and decision making—system-level goals or outcomes.
>
> Specifically, the initiative focuses on two overarching system-level outcomes to support and achieve the primary individual/client–centered outcome of improved health status. These system-level outcomes include: inclusive decision-making processes, and increased efficiency of the health-care system. To achieve these system-level outcomes, the program staff have worked to 1) establish an inclusive and accountable community decision-making process for fundamental health-care system reform; 2) achieve communitywide coverage through expansion of affordable insurance coverage and enhanced access to needed health-care services; and 3) develop a comprehensive, integrated delivery system elevating the roles of health promotion, disease prevention, and primary care, and integrating medical, health, and human services. These key objectives and the activities associated with achieving them are linked directly to the system-level goals of inclusive decision making and increased efficiency of the health-care system.

However, program staff found that it was easy in the stress of day-to-day work pressures to lose sight of the fact that the activities they were involved in to achieve system-level outcomes were not ends in themselves, but critical means to achieving the key client-level outcome of *improved health status*. To address this issue, project leaders in one community developed an effective method to assist staff and stakeholders in keeping the connection between systems and client-centered outcomes at the forefront of their minds. This method entailed "listening" to the residents of the communities where they operated. Program staff interviewed nearly 10,000 residents to gather input on how to improve the health status of those living in that community. Staff then linked these evaluation results to the system-level outcomes and activities they were engaged in on a daily basis. In this way, they were able to articulate clear connections between what they were doing at the system level (improving decision-making processes and efficiency), and the ultimate goal of improving the health status of community residents.

Broader Family or Community Outcomes:
It is also important to think more broadly about what an individual-level outcome really means. Many programs are aimed at impacting families, neighborhoods, and in some cases, whole communities. Besides individual outcomes, you and your staff need to think through the family and community-level outcomes you are trying to achieve—both interim and long-term. For instance, family outcomes might include improved communication, increased parent-child-school interactions, keeping children safe from abuse. Community outcomes might include increased civic engagement and participation, decreased violence, shifts in authority and responsibility from traditional institutions to community-based agencies and community resident groups, or more intensive collaboration among community agencies and institutions.

Impacts on Organizations
In addition to a project's external outcomes, there will also be internal effects—both individual and institutional—which are important to understand and document. Many times these organizational outcomes are linked to how effectively the program can achieve individual client outcomes. They are also important to understand in order to improve program management and organizational effectiveness. Questions to consider in determining these outcomes include:

Impact on personnel:
How are the lives and career directions of project staff affected by the project? What new directions, career options, enhanced perceptions, or improved skills have the staff acquired?

Impact on the institution/organization:
How is the home institution impacted? Does the presence of a project create ripple effects in the organization, agency, school, or university housing it? Has the organization altered its mission or the direction of its activities or the clientele served as a result of funding? Are collaborations among institutions strengthened?

Developing and Implementing an Outcome Evaluation Process
As we described above, an important first step of any outcome evaluation process is to help program staff and key stakeholders think through the different levels of program outcomes, and understand the importance of starting with individual client/participant outcomes rather than program or systems goals.

Once program staff and stakeholders have an understanding of outcome evaluation and how it can be used, you and your evaluation team can address the following questions which will facilitate the development of an outcome evaluation process:

1. Who are you going to serve?
2. What outcomes are you trying to achieve for your target population?
3. How will you measure whether you've achieved these outcomes?
4. What data will you collect and how will you collect it?
5. How will you use the results?
6. What are your performance targets?

(Framework based on Patton's work with Kellogg described in *Utilization-Focused Evaluation,* 1997 Edition)

1. Who are you going to serve? Before you and your program staff can determine individual client-level outcomes, you need to specify your target population. Who are you going to serve? Who are your clients/participants? It is important to be as specific as possible here. You may determine that you are serving several subgroups within a particular target population. For instance, a program serving women in poverty may find they need to break this into two distinct subgroups with different needs—women in corrections and women on welfare.

If your program serves families, you may have an outcome statement for the family as a unit, along with separate outcomes for parents and children. Here again, you would need to list several subgroups of participants.

2. What outcomes are you trying to achieve? Once you have determined who you are serving, you can begin to develop outcome statements. What specific changes do you expect in your clients' lives? Again, these changes might include changes

in behavior, knowledge, skills, status, level of functioning, etc. The key is to develop clear statements that directly relate to changes in individual lives.

3. How will you measure outcomes? In order to determine how effective a program is, you will need to have some idea of how well outcomes are being achieved. To do this, you will need ways to measure changes the program is supposed to effect. This is another place where program staff and stakeholders can lose sight of individual participant outcomes and begin to focus exclusively on the criteria or indicators for measuring these outcomes.

Outcomes and indicators are often confused as one and the same, when they are actually distinct concepts. Indicators are measurable approximations of the outcomes you are attempting to achieve. For example, self-esteem, in and of itself, is a difficult concept to measure. A score on the Coopersmith self-esteem test is an indicator of a person's self-esteem level. Yet, it is important to remember that the individual client-level outcome is not to increase participants' scores on the Coopersmith, but to increase self-esteem. The Coopersmith test simply becomes one way to measure self-esteem.

This program might also have constructed teacher assessments of a child's self-esteem to be administered quarterly. Here the indicator has changed from a standardized, norm-referenced test to a more open-ended, qualitative assessment of self-esteem; however, the outcome remains the same—increased self-esteem.

4. What data will you collect and how will you collect it? The indicators you select for each outcome will depend on your evaluation team's philosophical perspective about what is the most accurate measure of your stated outcomes; the resources available for data collection (some indicators are time- and labor-intensive to administer and interpret, e.g., student portfolios vs. standardized achievement tests); and privacy issues and how intrusive the data collection methods are. Your team should also consider the current state of the measurement field, reviewing the indicators, if any, that currently exist for the specific outcomes you are attempting to measure. To date, little work has been completed to establish clear, agreed-upon measures for the less concrete outcomes attempted by comprehensive, community-based initiatives (e.g., changes in community power structures; increased community participation, leadership development and community building) (Connell, Kubisch, Schorr, Weiss, 1995).

Another common problem is that all too often programs start with this step—by determining what can be measured. Program staff may then attempt to achieve only those outcomes which they know how to measure or which are relatively easy to measure. Since the field of measurement of human functioning will never be able to provide an accurate and reliable measure for every outcome

(particularly more complex human feelings and states), and since program staff and stakeholders often are knowledgeable about only a subset of existing indicators, starting with measures is likely to limit the potential for the program by excluding critical outcomes. The Kellogg Foundation believes it is important to start with the overall goals and outcomes of the program, and then determine how to go about measuring these outcomes. From our perspective, *it is better to have meaningful outcomes which are difficult to measure than to have easily measurable outcomes which are not related to the core of a program that will make a difference in the lives of those served.*

5. How will you use results? Ultimately, you want to ensure that the findings from your outcome evaluation process are useful. We suggest that you and your evaluation team discuss how you will use the results of the evaluation process from the beginning. Before you have even finalized data collection strategies, think through how you will use different outcome data and what specific actions you might take, depending on the findings. This will increase the likelihood that you will focus on the critical outcomes, select the most accurate and meaningful indicators, collect the most appropriate data, and analyze and interpret the data in the most meaningful ways. In addition, it will increase the likelihood that you and your staff will act on what you find, because you understood from the beginning what you were collecting and why you were collecting it.

6. What are your performance targets? Think of performance targets as benchmarks or progress indicators that specify the level of outcome attainment you expect or hope for (e.g., the percentage of participants enrolled in postsecondary education; how many grade-level increases in reading ability). Setting meaningful performance targets provides staff and stakeholders with benchmarks to document progress toward achieving program outcomes. These benchmarks help clarify and provide specificity about where you are headed and whether you are succeeding.

It is often best to set performance targets based on past performance. Therefore, you may want to wait until you have some baseline outcome data before determining performance targets. However, if you do not have the luxury of waiting to collect baseline data, you can set initial performance targets based on levels attained in comparable or related programs.

Measuring the Impacts of System Change and Comprehensive Community-Based Initiatives

As discussed previously, we need to think differently about evaluating the impacts of more complex system change and comprehensive community initiatives. In

these initiatives, implementation is difficult and long, and requires a collaborative, evolutionary, flexible approach. We may not see ultimate outcomes for many years, and many of the desired outcomes are difficult to measure using traditional quantitative methodologies. And yet, these initiatives hold great promise for really making a difference in our communities.

When evaluating these initiatives, then, we need to use innovative methods, such as participatory and theory-based evaluation, to learn as much as we can about how and why these programs work. By working together to develop the key interim outcomes, we will be able to document better the progress of these initiatives, and to understand better how they lead to the desired long-term outcomes.

There are two categories of interim outcomes you should think about measuring. The first includes interim outcomes associated directly with your target population. For example, interim outcomes associated with the long-term outcome of getting off public assistance might include leaving abusive relationships or conquering a drug problem.

The second category of interim outcomes includes changes in the project's or community's capacity to achieve the long-term desired outcomes (Schorr and Kubisch, 1995). For a project designed to increase the number of students going to college, important interim outcomes might be the implementation of a new professional development program to educate guidance counselors and teachers about how to encourage and prepare students for college; increased student access to financial aid and scholarship information; or an expansion in the number and type of summer and after-school academic enrichment opportunities for students.

Measuring Impacts Through the Use of a Program Logic Model

One effective method for charting progress toward interim and long-term outcomes is through the development and use of a program logic model. As we discussed earlier, a program logic model is a picture of how your program works—the theory and assumptions underlying the program. A program logic model links outcomes (both short- and long-term) with program activities/processes and the theoretical assumptions/principles of the program. This model provides a roadmap of your program, highlighting how it is expected to work, what activities need to come before others, and how desired outcomes are achieved.

There are multiple benefits to the development and use of a program logic model. First, there are *program design benefits*. By utilizing a program logic

model as part of the evaluation process, staff will be able to stay focused better on outcomes; connect interim outcomes to long-term outcomes; link activities and processes to desired outcomes; and keep underlying program assumptions at the forefront of their minds. In short, the process of creating a program logic model will clarify your thinking about the program, how it was originally intended to work, and what adaptations may need to be made once the program is operational.

Second, the program logic model provides a powerful base from which to conduct *ongoing evaluation of the program*. It spells out how the program produces desired outcomes. In this way, you can decide more systematically which pieces of the program to study in determining whether or not your assumptions were correct. A program logic model helps focus the evaluation on measuring each set of events in the model to see what happens, what works, what doesn't work, and for whom. You and your evaluation team will be able to discover where the model breaks down or where it is failing to perform as originally conceptualized.

As we discussed, logic model or theory-based evaluation is also an effective approach for evaluating complex initiatives with intangible outcomes (such as increased community participation) or long-term outcomes that will not be achieved for several years. A program logic model lays out the interim outcomes and the more measurable outcomes on the way to long-term and intangible outcomes. As a result, it provides an effective way to chart the progress of more complex initiatives and make improvements along the way based on new information.

Finally, there is value in the *process of developing a logic model.* The process is an iterative one that requires stakeholders to work together to clarify the underlying rationale for the program and the conditions under which success is most likely to be achieved. Gaps in activities, expected outcomes, and theoretical assumptions can be identified, resulting in changes being made based on consensus-building and a logical process rather than on personalities, politics, or ideology. The clarity of thinking that occurs from the process of building the model becomes an important part of the overall success of the program. The model itself provides a focal point for discussion. It can be used to explain the program to others and to create a sense of ownership among the stakeholders.

Types of Program Logic Models:

Although logic models come in many shapes and sizes, three types of models seem to be the most useful. One type is an **outcomes model.** This type displays the interrelationships of goals and objectives. The emphasis is on short-term objectives as a way to achieve long-term goals. An outcomes logic model might

be appropriate for program initiatives aimed at achieving longer-term or intangible, hard-to-measure outcomes. By creating a logic model that makes the connections between short-term, intermediate and long-term outcomes, staff will be able better to evaluate progress and program successes, and locate gaps and weaknesses in program operations. See Figure 1, the Community Health Partnership Program Logic Model, for an example of this type.

Another type of logic model is an **activities model.** This type links the various activities together in a manner that indicates the process of program implementation. Certain activities need to be in place before other activities can occur. An activities logic model is appropriate for complex initiatives which involve many layers of activities and inter-institutional partnerships. In these cases, every stakeholder needs to have the big picture of how the activities and processes pull together into a cohesive whole to achieve desired outcomes. It also provides an effective means to document and benchmark progress as part of the evaluation process. Which activities have been completed? Where did the program face barriers? How successfully were activities completed? What additional activities and processes were discovered along the way that are critical to program success? An example of this type of program logic model can be seen in Figure 2, the Calhoun County Health Improvement Program Logic Model.

The third type of logic model is the **theory model.** This model links theoretical constructs together to explain the underlying assumptions of the program. This model is also particularly appropriate for complex, multi-faceted initiatives aimed at impacting multiple target populations (e.g., multiple members of a family, whole communities, multiple institutions or community organizations within a community, etc.). At the same time, a theory logic model is also effective for a simpler program because of its ability to describe why the program is expected to work as it does. See Figure 3, the Conceptual Model of Family Support, for an example of this type.

Oftentimes, program staff will find that they will need to combine two or three of these program logic models. See Figure 4, the Human Resource Management for Information Systems Strategy Network, for an example of this hybrid.

Figure 1
Outcomes Model
Community Health Partnership Program Logic Model

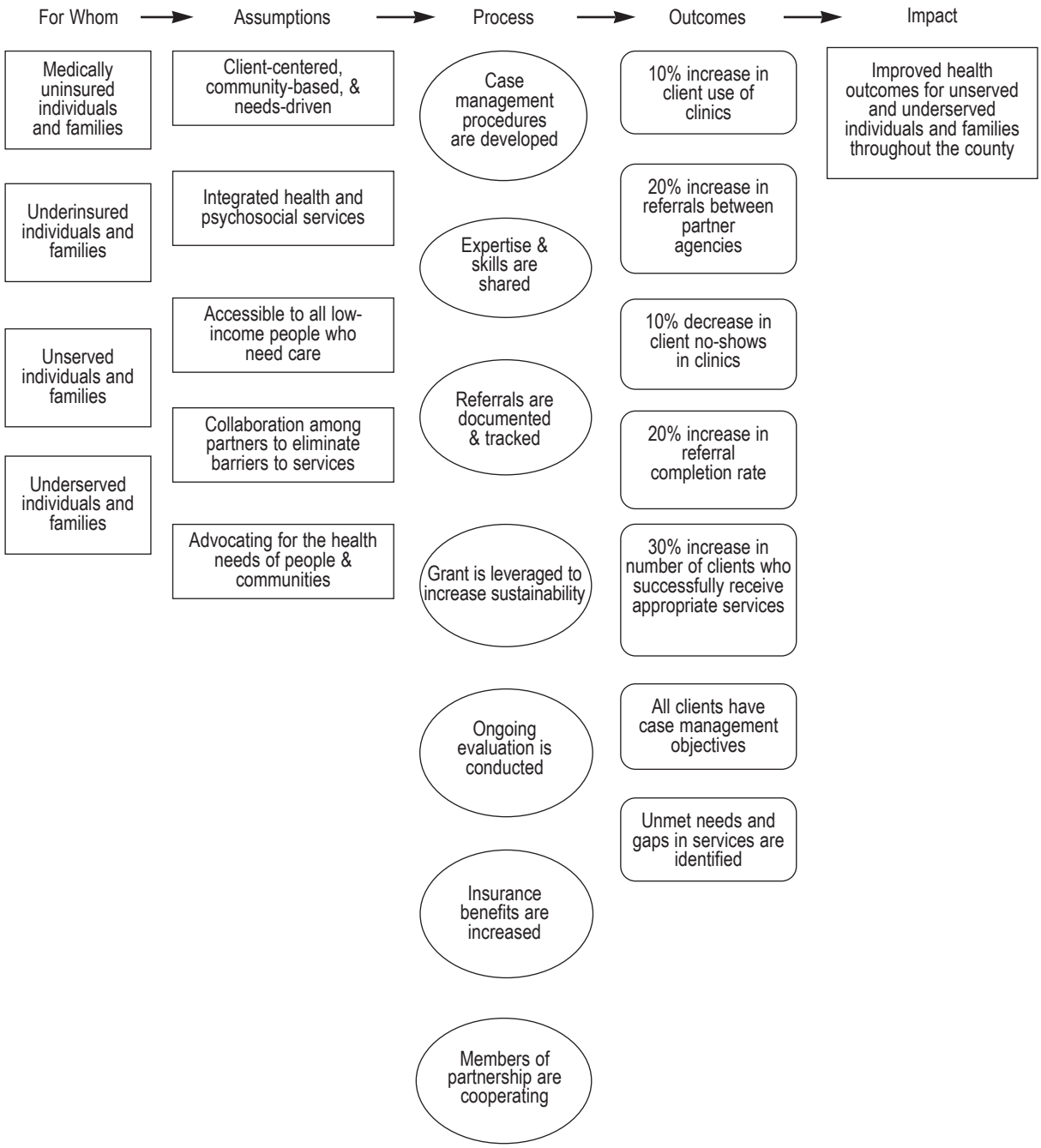


Figure 2
Activities Model
Calhoun County Health Improvement Program Logic Model

A Required Input/ Resources (Ongoing)	B Planning Phase Activities (1994-96)	C Planning Phase Outcomes (1994-96)	D1 Short-Term Implementation Phase Activities (1996-97)
1. Community leaders committed to the development of a shared vision for improved health status county-wide (A-1-4). 2. Broad base of citizens committed to systemic reform of county health care service delivery (A-1). 3. Philosophy of continuous program improvement through shared, data driven decision-making and capacity building (A-1-5). 4. Neutral group to catalyze and integrate the reform dialogue into required action. 5. Neutral fiscal agent/convener and community financial support sufficient to sustain activities post-grant (A-2). 6. Technical expertise on insurance, health care, community advocacy, and telecommunication issues (A-1-5). 7. Strategic planning, management, marketing evaluation, and public relations expertise (A-3-6).	1. Establishing community decision-making process (B-2). 2. Establishing administrative structure for program (B-3). 3. Establishing workgroups to gather community input and recommend improvement plan (B-4). 4. Conduct community meetings to gain feedback/sanction for vision and planning (B-1). 5. Develop strategic plan to achieve community derived vision for improved health status (B-5). 6. Design and implement preliminary needs assessments, communication and outreach activities (B-5).	1. Linkages formed among existing community leaders/ stakeholders (C-1). 2. Structure and staff for implementation established (C-2). 3. Implementation teams formed (C-3). 4. Community vision for systemic health care reform drafted and approved (C-6). 5. Policy changes—health plans, data exchange, service integration, Medicaid-identified to drive planning and aid implementation phase (C-5). 6. Community funding provided to support telecommunications network and other activities (C-3,5,6). 7. Public support evident for community derived vision (C-4, 6).	1. Development, pilot testing, and promotion of shared decision-making model. (D-5). 2. Build stakeholder capacity to influence local policy through recruitment and education (D-6). 3. Consumers, payers, and providers sought and encouraged to serve together on CCHIP convened boards/ working committees to achieve common goals (D-1,2,6). 4. Model development-community access/converge issues identified by research (D-3). 5. Public relations, marketing, and consumer advocacy programs developed to support enrollment strategy (D-4-6). 6. Development of exchange protocols that support expansion of shared network (D-3,4,6). 7. Development of training and support services to facilitate service delivery and growth (D3,4). 8. Contract with CCHIP to implement ongoing community health assessment (D-3,4,6). 9. Support provided for community leadership of health service improvement projects (D-1,2). 10. Development of training and evaluation activities to build capacity of health promotion organizations (D-2,6).

Continued on next page

Figure 2 (continued)
Activities Model

Calhoun County Health Improvement Program Logic Model

D2 Short-Term Implementation Phase Outcomes (1996-97)	E Intermediate-Term Implementation Phase Outcomes (1997-98)	F Long-Range Outcomes (1998-99)	G Desired Social Change
1. Shared decision-making model disseminated to local health care organizations (E-1, 2). 2. Improved capacity of Membership Organization to influence public policy (E1-3). 3. Improved communication and inter-organizational relations attributed to project activity (E-2). 4. Strategic planning assists stakeholders to achieve their shared vision—improved health status in Calhoun County (E-3, 4). 5. Third party administrator contract solicitation/ award guided by Health Plan Purchasing Alliance board criteria (E-4). 6. Healthplan contracts solicited by the Health Plan Purchasing Alliance Board (E-4). 7. Information exchange protocols and techno-logical/administrative infrastructure have the capacity to support service delivery (E-4). 8. Training and support contribute to Health Information Network system expansion (E-4). 9. Community health assessment data used to inform ongoing community health care decision making (E-4). 10. "811" primary care management and referral operational (E-4). 11. Increased local capacity to integrate health services (E-4). 12. Neighborhood health status improvement projects operational and supported by the community (E-4).	1. Local health care organizations increase use of shared decision making. 2. Research based community advocacy and influence molds public policy to impact community health status. 3. Payers and providers progressing toward coordination of resources and improved dispute resolution. 4. Improved access/ coverage for the under and uninsured in the community. 5. Increased number of health plan contracts secured. 6. Decentralization of medical records. 7. Health Information Network provides leverage for health care improvement. 8. Infrastructure and resources for sustaining periodic community health assessment in place. 9. Increased integration of health care delivery systems. 10. Primary care providers active in research-based disease management program. 11. Increased access/ participation - health promotion and primary care. 12. Community organizations make substantive contributions and provide ongoing support for health and primary care promotion. 13. Reduction in incidence of targeted health behavior.	1. Inclusive, accountable community health decision-making process. 2. Community administrative process which supports local points for health data, policy, advocacy, dispute resolution, and resource coordination. 3. Community-wide coverage with access to affordable care within a community-defined basic health service plan with a strategy to include the under- and uninsured. 4. Community-based health information systems which include performance monitoring, quality and cost effectiveness measurement, accessible records, and consumer satisfaction. 5. Community health assessment - utilizes community health profiles and indicators of access, health status, system resource performance and risk. 6. Comprehensive integrated health delivery system that elevates the roles of health promotion, disease prevention, and primary care, integrates medical, health, and human service systems.	1. Inclusive decision making. 2. Increased efficiency of health care system. 3. Improved health status.

Figure 3
Theory Model
Conceptual Model of Family Support

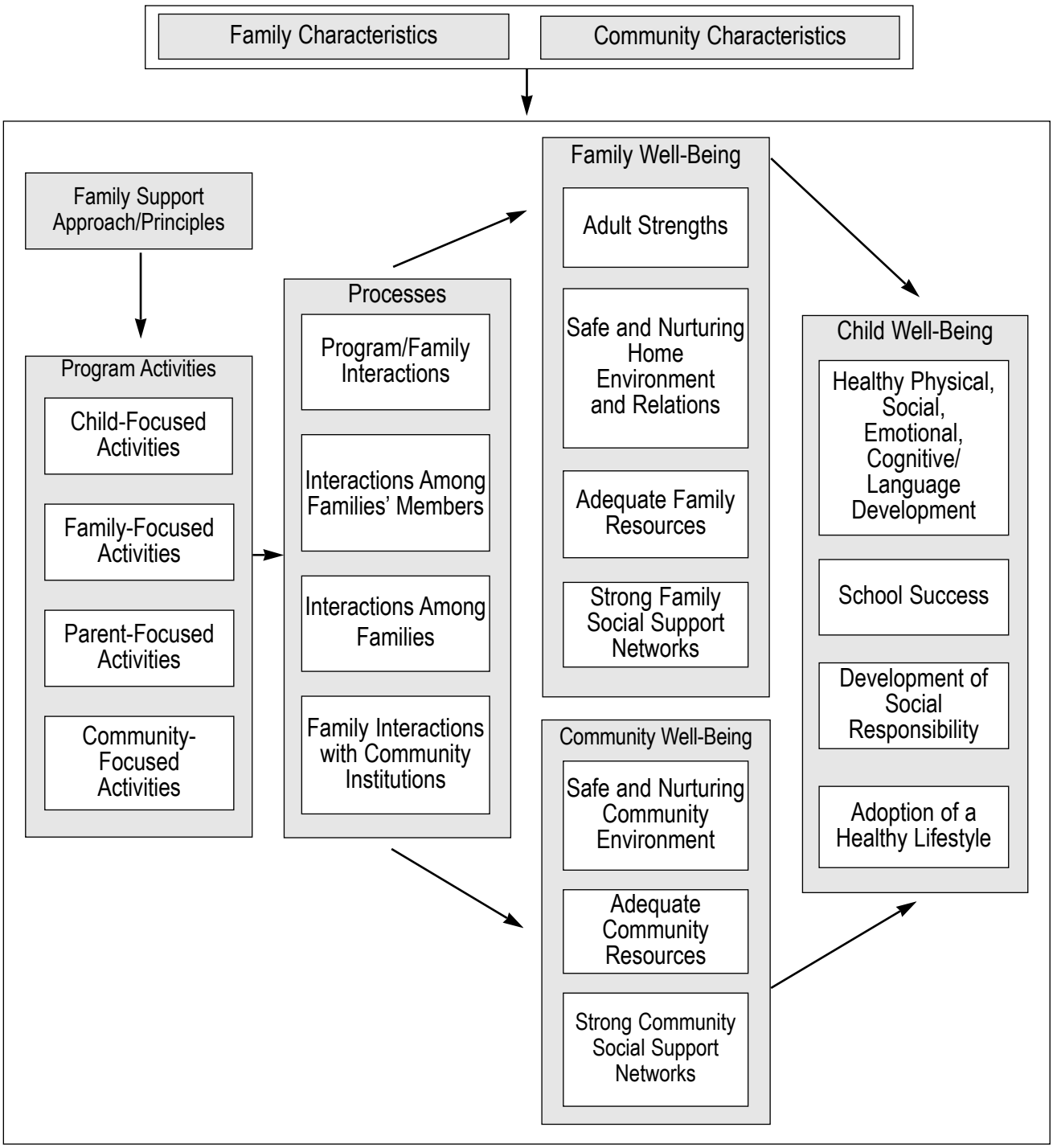
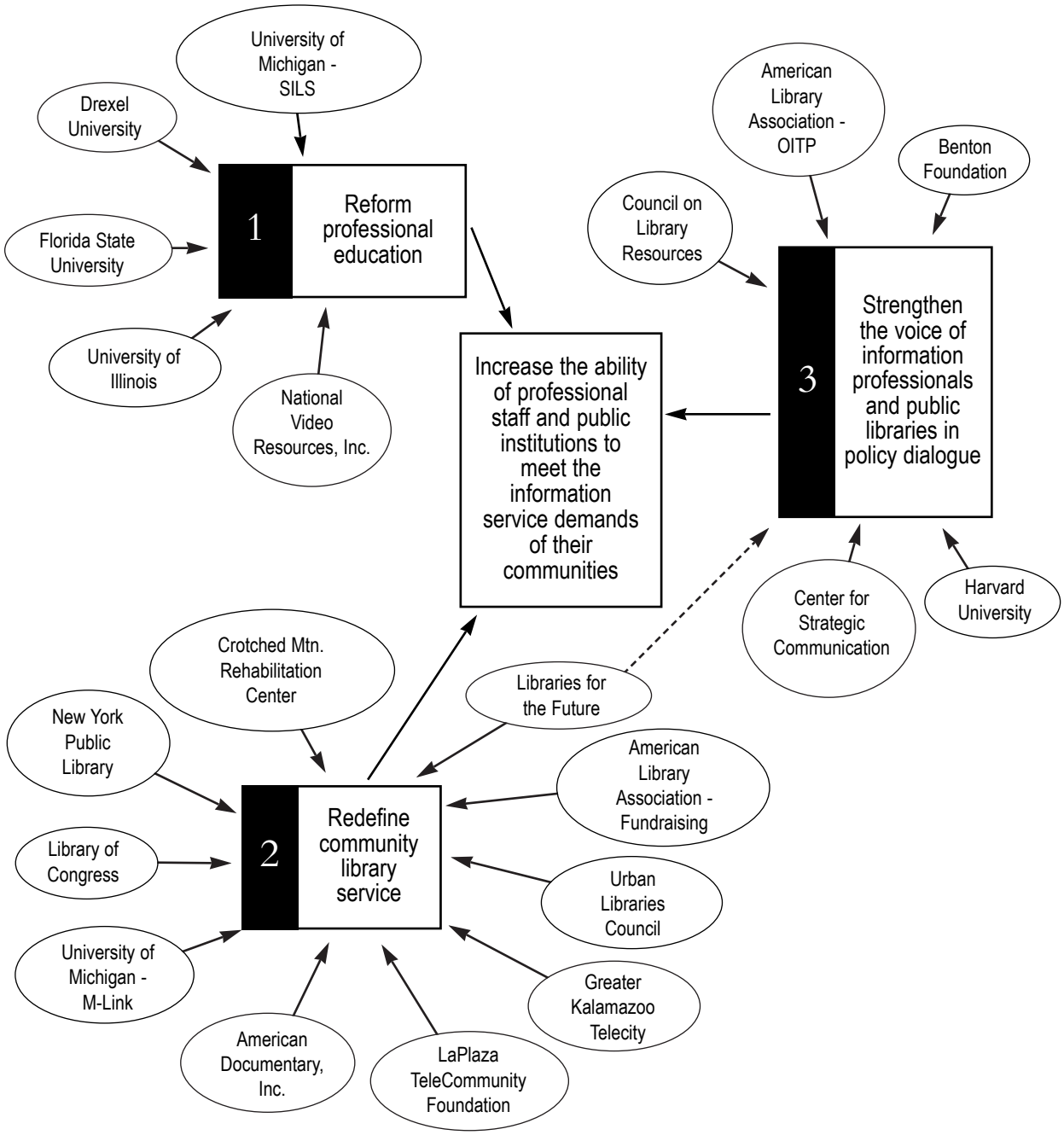


Figure 4
Hybrid/Combination Model
HRISM Strategy Network
1996



Building a Logic Model:
Logic models can be created in many different ways. The starting place could be the elements of an existing program which are then organized into their logical flow. For a new program that is in the planning phase, the starting place could be the mission and long-term goals of the program. The intermediate objectives that lead to those long-term goals are added to the model, followed by the short-term outcomes that will result from those intermediate objectives. An activity logic model can be built in the same way; long-range activities are linked to intermediate and short-range activities.

The key to building any model is to prepare a working draft that can be refined as the program develops. Most of a logic model's value is in the process of creating, validating, and then modifying the model. In fact, an effective logic model will be refined and changed many times throughout the evaluation process as staff and stakeholders learn more about the program, how and why it works, and how it is being operationalized. As you test different pieces of the model, you will discover which activities are working and which are not. You may also discover that some of your initial assumptions were wrong, resulting in necessary model revisions to adapt it to current realities. You will learn from the model and change your program accordingly; but you will also learn a great deal from putting the program into practice, which will inform the model and provide you with new benchmarks to measure. This iterative evaluation process will ultimately lead to continuous improvements in the program and in staff's and other stakeholders' understanding of the program and how and why it works.

Sustainability
A major outcome of importance to the Kellogg Foundation is sustainability. Activities associated with the project—if not the project itself—may be sustained through state and federal monies, funding from other foundations, private donors, or adoption by larger organizations. How successfully projects are able to develop a strategy for the transition from short-term funding sources to long-term funding may determine their future existence. These are important outcomes for potential funders.

In this context, attention is increasingly focused on the conditions surrounding promising social programs. There is a growing body of evidence that suggests a program's success over the long term is associated with the ability of key stakeholders to change the conditions within which programs operate, thereby creating an environment where programs can flourish. This ability to change the conditions within which the program operates has oftentimes been more

important to its ultimate success than the program's level of innovation. Given this, we need to pay attention to and document the social, political, cultural, and economic conditions which support or hinder a program's growth and sustainability, and identify effective strategies for creating supportive conditions and changing difficult or hostile environments.

Replication and Dissemination

In isolation, a project needs only to sustain itself; but to have an impact in a larger context, it needs to be understood well enough to replicate or to disseminate important lessons learned. Projects can do this through publishing journal articles; participating in networks of communities/projects grappling with similar issues; presenting information locally, regionally, or nationally; advising similar projects; or assisting in replicating the project in other communities. All of these activities are also project outcomes.

Impact on Policy

In addition, we have an interest in working with grantees to influence or shape policy at the local, state, or federal level. We understand that research and evaluation rarely affects policy directly; instead policy is influenced by a complex combination of facts, assumptions, ideology, and the personal interests and beliefs of policymakers. At the same time, it is critical to proactively design and utilize evaluation processes and results not only to improve practice, but also to improve and change policies at multiple levels. It is only through connecting policy and practice in meaningful ways that we can hope to make real and sustainable change in the lives of children, youth, and families in our communities.

Creating policy change may seem like an impossible task; however, this goal is often reached through collective action, networking, and collaboration. These activities should be documented, as well as any responses to policy-change efforts.

The following provides some ideas for thinking about policy impacts when designing and implementing evaluation:

Discuss policy goals up front. An effective way to begin the process of impacting policy is to build it into discussions about evaluation design and implementation from the beginning. First, ensure that everyone involved in the evaluation understands how policies and people in decision-making positions can make a difference in people's lives. From here, you can discuss the specific types of policies you are attempting to influence. Depending on the stage of your program development and program goals, these policies might range from operating

policies in your own organization, to policies of larger systems within which you work or with whom you partner (e.g., juvenile justice, local welfare office), to state and even national legislated policies. Continue discussions about policy impacts by asking questions such as: What levels and types of policies are we hoping to change? Who are we attempting to influence? What do they need to hear? Think in advance about what their concerns are likely to be and build these questions into your evaluation design so you can address these issues up front.

Finally, incorporate these policy goals into your evaluation design. In many cases, the types of programmatic and organizational evaluation questions you are attempting to address will have important policy aspects to them; so it will not always mean an additional set of questions—perhaps just additional aspects of existing evaluation questions.

Think about who your audience is. When communicating findings and insights, think carefully about who your audience is. How will they respond to what you have learned? What are their concerns likely to be? What questions will they ask? Be ready to respond to these concerns and questions. Remember that most policy and decision makers want simple and direct information which everyone can understand. They want to understand not only what the problem is but also how to improve people's lives. Stay away from abstract research jargon and explain what you have learned about how and why your program works.

No matter who your audience, present evaluation findings and lessons in compelling ways. This does not mean skewing data and misusing evaluation results to create falsely positive marketing and public relations materials. There are many far more compelling ways to present the findings of an evaluation than with traditional research reports full of technical jargon and research-defined categories and criteria. Present real-life stories and real issues. Your program and the people served should be recognizable and well understood in any report or presentation you give.

Communicate what you know to community members. An important part of policy change is to get information out to the general public. Public attitudes and beliefs have a profound impact on policy at all levels. Educating community members about relevant issues—particularly in this technological age when they are barraged with so much shallow and incomplete information—will create potentially powerful advocates and partners to help shape policy.

Program participants and graduates are another important group with whom to share evaluation findings and lessons. In this way, you will be investing in their knowledge and skill development and helping to amplify their voices. Participants

are often the most powerful advocates for policy changes because they can speak from both their experience and from the lessons learned from evaluation processes.

Be proactive. Within the legal limits mandated by federal laws, don't be afraid to advocate based on what you are learning. Nonprofits can become very leery when people start using words like "advocacy." Many evaluators also balk at the idea of an advocacy role. However, there is nothing wrong with spreading the word about what you know when it is backed up by evidence (both quantitative and qualitative) garnered through sound evaluation processes. If evaluation is conducted in a thoughtful, reflective, and careful way, you will learn much and there will be much to tell. Tell it.

Be direct about what types of policies might be effective based on findings. Don't leave it up to policymakers and legislators. They will more likely choose effective long-term solutions, if evaluations demonstrate their worth.

Communicate about the evaluation process as well as the results. As previously discussed, human and social services practitioners are struggling to find more effective ways to evaluate complex social problems and the social programs designed to address them. There is much to be learned and shared about the evaluation process itself. Many policymakers and funders, as well as practitioners and researchers, need to be educated about the strengths and weaknesses of particular evaluation approaches. Often, programs find themselves pressured into certain types of evaluation designs because of funder and policymaker beliefs about how evaluation should be conducted (i.e., the traditional way). By increasing the knowledge base about the strengths and limitations of different evaluation philosophies and approaches, we will create a more supportive environment for trying out new approaches.

Summary

Conducting an outcome evaluation will help you determine how well your project is progressing in its effort to improve the well-being of your target population. In addition, it will support the continued existence of your project by helping current and potential funders and policymakers understand what your program has achieved and how.

In the end, the results of an outcome evaluation should not be used simply to justify the existence of a project; that is, to provide evidence that it worked. Instead, it should be viewed as a source of important information which can promote project development and growth. This points again to the importance of conducting outcome evaluation in combination with context and implementation evaluations.

Chapter 5

Planning and Implementing Project-Level Evaluation

Understanding the Kellogg Foundation's philosophy, expectations, and key evaluation components is only half of the formula for conducting effective project-level evaluation. This section provides a blueprint which, though not a comprehensive how-to manual, will take you through the process of planning and implementing project-level evaluation as an integrated part of your project.

Three things become clear from the literature on evaluation and our experiences with grantees across many human service areas. The first is that *the process is different for every community and every project. There is no one right way of doing evaluation.* Each project serves a different mix of clients, uses different service delivery approaches, defines different outcomes, is at a different phase of development, and faces different contextual issues. Therefore, the evaluation process that you and your staff develop will depend in large part on local conditions and circumstances.

The second is that *there are certain critical elements or action steps along the way that every project must address to ensure that effective evaluation strategies leading to real improvements are developed and implemented.*

The third is the continuity of the planning process. *The steps described below are not necessarily linear in practice.* Although there is some order to how projects initially address the steps, they are evolutionary in nature. Different aspects of each step will come into play throughout the development and implementation of your evaluation strategies.

The next section discusses the key issues to consider for each step of the blueprint, which we have structured into two phases: planning and implementation. However, it is important to note that all planning steps do not stop with the onset of implementation, but are active and ongoing throughout implementation. Similarly, every implementation step requires up-front planning. In addition, the descriptions of the action steps are not exhaustive. Each project must bring together relevant stakeholders to discuss how to tailor the blueprint to its particular situation and the questions it is asking.

Finally, although each action step is important on its own, they are all interwoven. It is the well-conceived packaging of these steps into your own "blueprint for action" that will result in an effective and useful evaluation.

Planning Steps: Preparing for an Evaluation

Planning Steps

1. Identifying Stakeholders and Establishing an Evaluation Team
2. Developing Evaluation Questions
3. Budgeting for an Evaluation
4. Selecting an Evaluator

Step 1: Identifying Stakeholders and Establishing an Evaluation Team

All evaluations have multiple stakeholders. A stakeholder is defined as any person or group who has an interest in the project being evaluated or in the results of the evaluation. Stakeholders include funders, project staff and administrators, project participants or customers, community leaders, collaborating agencies, and others with a direct, or even indirect, interest in program effectiveness.

For example, stakeholders of a school-based program created to encourage the development of interpersonal and conflict resolution skills of elementary students might include the program's developers, participating teachers, the school board, school administrators, parents, the participating children, taxpayers, funders, and yes, even the evaluators. It is important to remember that evaluators (whether internal or external) are stakeholders, and not neutral third parties, as we so often think. Evaluators have a vested interest in what they are doing and care about doing it well.

To ensure that you have gathered multiple perspectives about the salient issues, involve as many stakeholders as possible in initial evaluation discussions. Otherwise, the evaluation is likely to be designed based on the needs and interests of only a few stakeholders—usually the ones with the most power—and may miss other important questions and issues of stakeholders who are not included at the table.

Of course, involving every stakeholder may not be realistic. However, try to consult with representatives from as many stakeholder groups as possible when designing or redesigning the evaluation plan, and provide them with timely

results and feedback. We also encourage you to involve a manageable subset of stakeholder representatives in an evaluation team or task force. This team should come together, face-to-face if possible, to make ongoing decisions about the evaluation. Continued use of this team throughout the evaluation process (not just at the beginning of evaluation design) may help reduce project staff's concerns about evaluation and increase the amount and reliability of information collected. It will also increase the likelihood that recommendations will be accepted and implemented.

Although this step may be time-consuming and fraught with the potential for conflict, it is one well worth the time and effort. Involving many stakeholders will help ensure that the evaluation process goes more smoothly: more people are invested and willing to work hard to get the necessary information; project staff concerns about evaluation are reduced; the information gathered is more reliable and comes from different perspectives, thus forcing the team to think through the meaning of contradictory information; and the recommendations are likely to be accepted by a broader constituency and implemented more fully and with less resistance.

> ***Example:*** Program staff of a successful long-term initiative focused on heightening public awareness about groundwater quality and drinking water issues created an evaluation team consisting of the project director, key staff members, and the local evaluator. Early in the process of developing an evaluation plan, the team realized that it needed information and input from additional "outside" stakeholders, particularly representatives from local governments, such as staff from the local utilities departments, building departments, planning commissioners, as well as key elected officials. These stakeholders, although not directly involved in the implementation of the project, were critical players in terms of influencing policy related to groundwater quality, as well as increasing awareness and problem solving with the community around decision making that might affect groundwater drinking quality.
>
> These stakeholders also provided a unique perspective to the team. They were going to be the ones most immediately affected by project staff's work, and were the ones best able to work with project staff to test questions and determine strategic action steps. In addition, the evaluation plan focused primarily on gathering information from these outside stakeholders; therefore, representatives from these groups needed to be a part of the discussions regarding data collection processes and structures. What was the best way to reach local government representatives?

Initially, staff decided to expand the primary evaluation team to include representatives from these additional stakeholder groups. However, it quickly became apparent that including everyone would make the evaluation team too large to operate effectively. In addition, calls to these potential representatives revealed another problem. Although many of the stakeholders contacted were interested in participating and providing their input, they were concerned when they learned about the level of effort and time commitment that would be required of them, given their already busy schedules. Being public officials, most of them had many roles to fill, including multiple committee appointments and other meetings, which went beyond their regular work hours. It did not seem feasible that these stakeholders could manage biweekly, or even monthly, evaluation team meetings.

However, instead of giving up and foregoing the important input from these stakeholders (as is often the case with project-level evaluations that involve multiple "outside" stakeholders), project staff decided to create a second ad hoc team made up of approximately 20 representatives from these stakeholder groups. This team was brought together at certain critical points in the process to provide feedback and input to the primary evaluation team. Specifically, they were brought together two to three times per year for roundtable discussions around particular evaluation topics, and to provide input into next steps for the program and its evaluation. An added benefit of these roundtables was that local representatives from multiple communities were able to problem solve together, learn from one another, and create a network of peers around groundwater issues—strengthening the program itself, as well as the evaluation component.

In addition, the primary evaluation team called on five representatives from these outside stakeholder groups on a more frequent basis for input into evaluation and programmatic questions and issues. In this way, the project was able to benefit from input from a wider variety of perspectives, while making participation in the evaluation a manageable process for all those involved.

Things to Remember . . .

- Gathering input from multiple stakeholders helps you remain aware of the many levels of interest related to the project. You and your evaluation team will be better prepared to counteract pressure from particular stakeholders for quick fixes or a rush to judgment when that is not what is best for the project.
- Stakeholders will have different, sometimes even contradictory, interests and views. They also hold different levels of power. Project directors have more power than staff. Legislators have more power than primary-grade students. Your funders have a particular kind of power. Ask yourself: Which stakeholders are not being heard in this process? Why not? Where can we build consensus and how can we prioritize the issues?
- Evaluators are stakeholders, too. What are their interests? How might this affect how the evaluation is designed, which questions are focused on, and what interpretations are made?

Step 2: Developing Evaluation Questions

Drafting an evaluation plan will most likely require numerous meetings with the evaluation team and other stakeholders. One of the first steps the team should work through is setting the goals of the evaluation. The main concern at this stage, and perhaps the biggest challenge, is to determine what questions need to be answered.

Again, questions will depend on the phase of project development, the particular local circumstances, and the ultimate purpose of the evaluation. Critical evaluation questions to address over the life of a project include, but are not limited to:

1. What do you want your project to accomplish?
2. How will you know if you have accomplished your goals?
3. What activities will your project undertake to accomplish your goals?
4. What factors might help or hinder your ability to accomplish your goals?
5. What will you want to tell others who are interested in your project?

Seek input on critical questions from a variety of sources. Some potential sources for this initial question formation period include:

Commitment Letter—The Kellogg Foundation always sends a commitment letter when a decision has been made to fund a project. This letter usually contains a list of evaluation questions the Foundation would like the project to address. Project evaluation should not be limited to answering just these questions, however.

Project Director—A director can be an invaluable source of information because he or she has probably been involved in project conceptualization, proposal development, and project design and implementation, and is therefore likely to have an overall grasp of the venture.

Project Staff/Volunteers—Staff members and volunteers may suggest unique evaluation questions because they are involved in the day-to-day operations of the project and have an inside perspective of the organization.

Project Clientele—Participants/consumers offer crucial perspectives for the evaluation team because they are directly affected by project services. They have insights into the project that no other source is likely to have.

Board of Directors/Advisory Boards/Other Project Leadership—These groups often have a stake in the project and may identify issues they want addressed in the evaluation process. They may request that certain questions be answered in order to help them make decisions.

Community Leaders—Community leaders in business, social services, and government can speak to issues underlying the conditions of the target population. Because of their extensive involvement in the community, they often are invaluable sources of information.

Collaborating Organizations—Organizations and agencies that are collaborating with the grantee should always be involved in formulating evaluation questions.

Project Proposal and Other Documents—The project proposal, Foundation correspondence, project objectives and activities, minutes of board and advisory group meetings, and other documents may be used to formulate relevant evaluation questions.

Content-Relevant Literature and Expert Consultants—Relevant literature and discussion with other professionals in the field can be potential sources of information for evaluation teams.

Similar Programs—Evaluation questions can also be obtained from directors and staff of other projects, especially when these projects are similar to yours.

Obviously you may not have access to all of these sources, nor will you be able to explore all of the questions identified. Therefore, the next step is to have a representative sample of stakeholders prioritize the list of potential questions to determine which will be explored. One way to prioritize questions is to list them clearly on a large sheet of paper so that they can be discussed easily in a meeting. Using memos is another way to share the potential questions that your team is considering.

Keep in mind that questions may need to be made more specific in order to be answerable. For example, in order to address the broad question, "What has this project done for participants?" you may need to ask several more specific questions, such as:

- Has the project improved attendance at clinic visits?
- Has the project had an impact on the emotional needs of participants?
- Has the project decreased the number of teenage pregnancies in our clientele in comparison to other teenagers in the area?

An effective way to narrow the possible field of evaluation questions is through the development of a program logic model. As we have discussed throughout this handbook, a program logic model describes how your program works. Often, once you have built consensus on a program logic model, you will find that the model provides you and your evaluation team with a focus for your evaluation. The model helps clarify which variables are critical to achieving desired outcomes. Given the vast array of questions you may want to answer about your program, a program logic model helps narrow the field in a systematic way by highlighting the connections between program components and outcomes, as well as the assumptions underlying the program. You will be better able to address questions such as: How is the program supposed to work? Where do the assumptions in the model hold and where do they break down? Where are the gaps or unrealistic assumptions in the model? Which pieces of the model seem to be yielding the strongest impacts or relationships to one another? Which pieces of the model are not being operationalized in practice? Are there key assumptions that have not been embedded in the program that should be?

By organizing evaluation questions based on your program's logic model, you will be better able to determine which questions to target in an evaluation. You

will also be better able to use what you find out to improve the program by revising the model and then developing an action plan to operationalize these changes in practice.

Things to Remember . . .

- The particular philosophy of evaluation/research that you and your evaluation team members espouse will influence the questions you ask. Ask yourself and team members why you are asking the questions you are asking and what you might be missing.
- Different stakeholders will have different questions. Don't rely on one or two people (external evaluator or funder) to determine questions. Seek input from as many perspectives as possible to get a full picture before deciding on questions.
- There are many important questions to address. Stay focused on the primary purpose for your evaluation activities at a certain point in time and then work to prioritize which are the critical questions to address. Since evaluation will become an ongoing part of project management and delivery, you can periodically revisit your evaluation goals and questions and revise them as necessary.
- Examine the values embedded in the questions being asked. Whose values are they? How do other stakeholders, particularly project participants, think and feel about this set of values? Are there different or better questions the evaluation team members and other stakeholders could build consensus around?

Step 3: Budgeting for an Evaluation

Conducting an evaluation requires an organization to invest valuable resources, including time and money. The benefits of a well-planned, carefully conducted evaluation outweigh its costs; therefore, the Kellogg Foundation expects that a portion of your budget will be designated for evaluation. Generally, an evaluation costs between 5 and 7 percent of a project's total budget. If, in your initial proposal, you underestimated your evaluation costs, you should request additional funds.

Although it is likely you will need to revise specific pieces of your evaluation budget as you fill in the details of your design and begin implementation, you should consider your evaluation budget as part of the up-front planning step.

Worthen and Sanders (1987) provide a useful framework for developing an evaluation budget (see Worksheet B). You should modify this framework as appropriate for your organization, or if you have other expenses that are not listed. The categories of their framework include:

1. Evaluation staff salary and benefits—The amount of time staff members must spend on evaluation and the level of expertise needed to perform particular evaluation tasks will affect costs.
2. Consultants—If your staff needs assistance in conducting the evaluation, you will need to contract with external consultants. These consultants can provide special expertise and/or different perspectives throughout the process of evaluation.
3. Travel—Travel expenses for staff and/or evaluators vary from project to project. Projects located far from their evaluators or projects with multiple sites in different parts of the country may need a large travel budget. In addition, all projects need to budget for transportation costs to Foundation evaluation conferences.
4. Communications—You will have to budget for communication costs, such as postage, telephone calls, etc.
5. Printing and duplication—These costs cover preparation of data-collection instruments, reports, and any other documents.
6. Printed materials—This category includes the costs of acquiring data-collection instruments and library materials.
7. Supplies and equipment—This category covers the costs of specific supplies and equipment (e.g., computers, packaged software) that must be purchased or rented for the evaluation.

Worksheet B

Planning an Evaluation Budget

	Year 1	Year 2	Year 3	Year 4
1. Evaluation Staff Salary and Benefits	$	$	$	$
2. Consultants	$	$	$	$
3. Travel	$	$	$	$
4. Communications (postage, telephone calls, etc.)	$	$	$	$
5. Printing and Duplication	$	$	$	$
6. Printed Materials	$	$	$	$
7. Supplies and Equipment	$	$	$	$
8. Other	$	$	$	$
9. Other	$	$	$	$
10. Total	$	$	$	$

Things to Remember . . .

- It is likely you will need to make changes to your budget. Build in a mechanism for reviewing and revising your evaluation budget based on what happens during the course of the evaluation.
- Often evaluation budgets do not provide enough resources for the analysis and interpretation step. To ensure sufficient resources for analysis and interpretation, think about how to save time and money during the design, data-collection, and reporting phases.
- Qualitative evaluation studies based on interpretivist/constructivist assumptions can be very effective at getting inside the program and really understanding how and why it works. However, these studies often are more costly to implement, since they require significant time talking with and observing many people involved with the project. Think about ways to utilize project staff, volunteers, and residents and incorporate qualitative collection and analysis techniques into the day-to-day operations of the project.
- Consider developing an evaluation time budget, as well as a cost/resources budget. The time required for evaluation will vary depending on the questions you are attempting to answer, the human and financial resources you have available, as well as other external factors. It is important to think through timing issues to ensure that your evaluation is feasible and will provide you with accurate, reliable, and useful information. Many projects fail to budget enough time for an evaluation, and then find at the conclusion of the process that the evaluation was not as helpful or useful as originally expected. A time budget can go a long way to addressing these issues.

Step 4: Selecting an Evaluator

As noted earlier, we do not require that all projects have an independent evaluator. In fact, this handbook is written to encourage and aid grantees in conducting their own evaluations. Evaluation is a critical thinking process, and successful project directors and staff do not readily delegate their thinking to someone outside the project. Still, there are times when it is appropriate to utilize people with evaluation expertise.

Types of Evaluators

In general, there are three types of evaluators: external evaluators, internal evaluators, and internal evaluators with an external consultant. You must determine what type of evaluator would be most beneficial to your project.

External Evaluator:

External evaluators are contracted from an outside agency or organization to conduct the evaluation. These evaluators often are found at universities, colleges, hospitals, consulting firms, or within the home institution of the project. Because external evaluators maintain their positions with their organizations, they generally have access to more resources than internal evaluators (i.e., computer equipment, support staff, library materials, etc.). In addition, they may have broader evaluation expertise than internal evaluators, particularly if they specialize in program evaluation or have conducted extensive research on your target population. External evaluators may also bring a different perspective to the evaluation because they are not directly affiliated with your project. However, this lack of affiliation can be a drawback. External evaluators are not staff members; they may be detached from the daily operations of the project, and thus have limited knowledge of the project's needs and goals, as well as limited access to project activities.

Internal Evaluator:

A second option is to assign the responsibility for evaluation to a person already on staff or to hire an evaluator to join your project. This internal evaluator could serve as both an evaluator and a staff member with other responsibilities. Because an internal evaluator works within the project, he or she may be more familiar with the project and its staff and community members, have access to organizational resources, and have more opportunities for informal feedback with project stakeholders. However, an internal evaluator may lack the outside perspective and technical skills of an external evaluator.

When hiring an internal evaluator, keep in mind that university degrees in evaluation are not common; many people now working as evaluators have previously held managerial/administrative roles or conducted applied social research. Consider hiring those who do not label themselves as professional evaluators, but who have conducted evaluation tasks for similar projects.

Internal Evaluator with an External Consultant:

A final option combines the qualities of both evaluator types. An internal staff person conducts the evaluation, and an external consultant assists with the technical aspects of the evaluation and helps gather specialized information.

With this combination, the evaluation can provide an external viewpoint without losing the benefit of the internal evaluator's first-hand knowledge of the project.

The Evaluator's Role

Whether you decide on an external or internal evaluator or some combination of both, it is important to think through the evaluator's role. As the goals and practices of the field of program evaluation have diversified, so too have evaluators' roles and relationships with the programs they evaluate. (At the same time, it is important to note that the idea of multiple evaluator roles is a controversial one. Those operating within the traditional program evaluation tenets still view an evaluator's role as narrowly confined to judging the merit or worth of a program.)

From our view, the primary goals of evaluation are that stakeholders are engaged, active participants in the process and that the evaluation process and findings will be meaningful and useful to those ultimately responsible for improving and assessing the program. In the end, this means that there is no one way to do evaluation. Given that premise, the critical skills of an effective evaluator include the ability to listen, negotiate, bring together multiple perspectives, analyze the specific situation, and assist in developing a design with the evaluation team that will lead to the most useful and important information and final products.

With your staff and stakeholders, think through all of the potential evaluator roles and relationships and determine which configuration makes the most sense given your particular situation, the purpose of the evaluation, and the questions you are attempting to address.

One important role to think through is the relationship between the evaluator and primary stakeholders or the evaluation team. Questions to consider include: Should this relationship be distant or highly interactive? How much control should the evaluator have over the evaluation process as compared to the stakeholders/evaluation team? How actively involved should key staff and stakeholders be in the evaluation process?

Depending on the primary purpose of the evaluation and with whom the evaluator is working most closely (funders vs. program staff vs. program participants or community members), an evaluator might be considered a consultant for program improvement, a team member with evaluation expertise, a collaborator, an evaluation facilitator, an advocate for a cause, or a synthesizer. If the evaluation purpose is to determine the worth or merit of a

program, you might look for an evaluator with methodological expertise and experience. If the evaluation is focused on facilitating program improvements, you might look for someone who has a good understanding of the program and is reflective. If the primary goal of the evaluation is to design new programs based on what works, an effective evaluator would need to be a strong team player with analytical skills.

Experience tells us, however, that *the most important overall characteristics to look for in an evaluator are the ability to remain flexible and to problem solve.*

Figure 5 on the following page provides a summary of some of the special challenges programs may face and how these challenges affect an evaluator's role.

Figure 5

Challenges Requiring Special Evaluator Skills

Situation	Challenge	Special Evaluator Skills Needed
1. Highly controversial issue	Facilitating different points of view	Conflict-resolution skills
2. Highly visible program	Dealing with program publicity; reporting findings in a media-circus atmosphere	Public presentation skills Graphic skills Media-handling skills
3. Highly volatile program environment	Adapting to rapid changes in context, issues, and focus	Tolerance for ambiguity Rapid responsiveness Flexibility Quick learner
4. Cross-cultural or international program	Including different perspectives, values; being aware of cultural blinders and biases	Cross-cultural sensitivity Skilled in understanding and incorporating different perspectives
5. Team effort	Managing people	Identifying and using individual skills of team members; team-building skills
6. Evaluation attacked	Preserving credibility	Calm; able to stay focused on evidence and conclusions
7. Corrupt program	Resolving ethical issues/upholding standards	Integrity; clear ethical sense; honesty

Adapted from Patton, 1997, *Utilization-Focused Evaluation,* p. 131.

How to Find an Evaluator

Perhaps the greatest frustration grantees have in the area of evaluation is finding a qualified evaluator. Colleges and universities are often good sources, as are many private firms listed under "Management Consultants" in telephone directories. Your program director at the Kellogg Foundation can also suggest evaluators. The Foundation's Evaluation Unit maintains a resource bank of evaluation consultants.

Evaluator Qualifications

Before interviewing prospective evaluators, you should determine the qualifications you would like an evaluator to have. You may require, for example, someone with knowledge of the community or who has experience working with your target population. Others may desire an evaluator who has experience in a specific subject matter or in conducting evaluations of a certain type of project.

Worthen and Sanders (1987) suggest some basic qualifications that evaluators should possess, including formal training in evaluation, other educational experiences related to evaluation, a professional orientation (suited to the project's orientation), previous performance of evaluation tasks, and personal styles/characteristics that fit with your organization. The following page (see Worksheet C) provides more detail on these basic qualifications.

Worksheet C
Checklist for Selecting an Evaluator

	Evaluator appears to be: (Check one for each item)		
1. To what extent does the formal training of the potential evaluator qualify him/her to conduct evaluation studies? (Consider major or minor degree specializations; specific courses in evaluation methodology; whether the potential evaluator has conducted applied research in a human service setting, etc.)	Well Qualified	Not Well Qualified	Cannot Determine if Qualified
2. To what extent does the previous evaluation experience of the potential evaluator qualify him/her to conduct evaluation studies? (Consider items such as length of experience; relevance of experience.)			
3. To what extent is the professional orientation of the potential evaluator a good match for the evaluation approach required? (Consider items such as philosophical and methodological orientations.)	Acceptable Match	Unacceptable Match	Cannot Determine Match
4. To what extent does the previous performance of the potential evaluator qualify him/her to conduct evaluation studies for your project? What prior experience does she or he have in similar settings? (Look at work samples or contact references.)	Well Qualified	Not Well Qualified	Cannot Determine if Qualified
5. To what extent are the personal styles and characteristics of the potential evaluator acceptable? (Consider such items as honesty, character, interpersonal communication skills, personal mannerisms, ability to resolve conflicts, etc.)	Acceptable	Unacceptable	Cannot Determine Acceptability
Summary Based on the questions above, to what extent is the potential evaluator qualified and acceptable to conduct the evaluation?	Well Qualified and Acceptable	Not Well Qualified and/or Unacceptable	Cannot Determine if Qualified or Acceptable

When to Hire an Evaluator

While there is no "best" time to hire an evaluator, experience has shown that successful projects hire evaluators sooner rather than later. Ideally, evaluators can assist as early as the proposal-writing stage. If this is not possible, you should try to hire your evaluator before services are provided, or at the latest, during the first few months of the project. *Never* wait until your first annual report to the Foundation is due.

Contractual Arrangements

Contractual arrangements for hiring an evaluator vary from project to project. In many cases, the evaluator is already an employee of the project. Other projects decide to hire an evaluator as a part-time or full-time employee, and provide a salary and benefits comparable to other employees in the organization.

It is not uncommon for our grantees to develop a contract with an external consultant who works either as an individual or as part of an organization. Contracting with a consultant is easier for projects to administer since they do not have to provide benefits or withhold income taxes. However, contracts also allow for the least control over the direction of the evaluation.

Many grantees develop a Request for Proposals (RFP) and solicit written bids from several potential consultants. These bids, as well as the consultants' qualifications, are then considered by project staff when selecting an evaluator. A written, signed contract should be executed before the consultant begins working for your project. Your organization might have very specific rules about the bidding and contract process, including who is authorized to sign contracts on your organization's behalf. Be sure to familiarize yourself with your organization's policy before developing an RFP.

Evaluation consultants can be paid in a variety of ways; this is something you need to negotiate with your consultant before a contract is signed. Small consulting contracts are sometimes paid in one lump sum at the end of a contract or when the final evaluation report is submitted. Larger contracts are often paid in monthly installments upon the consultant's submission of a detailed time log.

Working as an Evaluation Team

Remember that when developing and implementing an evaluation, it is helpful to work as a team with staff, relevant stakeholders, and the evaluator. Through your combined efforts and expertise, a well-planned evaluation can emerge, and it will not be a mysterious process that only the evaluator understands. One way to ensure the success of this team effort is to communicate with each other in clear

and straightforward language (no jargon), in an atmosphere that is open to new ideas and respectful of alternative viewpoints. Finally, you should maintain contact on a regular basis and develop a system for settling differences and grievances. While you do not have to agree on all or most evaluation matters, open discussions and a feeling of camaraderie should exist.

Example 1: A program designed to provide educational services to families and children in an economically disadvantaged urban community was piloted through a multi-institutional partnership, including several African-American churches, health centers, and educational providers. In selecting an evaluator, staff from the partner agencies were concerned about the characteristics and background of the person they would eventually hire. They were operating their program within a primarily African-American community, and felt that African Americans and their communities had been exploited within the traditional research communities. Therefore, they were skeptical of traditional avenues for finding evaluators, such as universities, research organizations, and private consulting firms.

In addition, their program was centered within the African-American church and was focused on the importance of spirituality to a fulfilling and self-sufficient life. To this end, they wanted an evaluator who was sensitive to the nuances and meaning of African-American spirituality and the church. Without this familiarity, program staff felt the evaluator would miss critical parts of the program's story or misunderstand aspects of the program or its impact. In addition, they wanted an evaluator with the methodological expertise to help them determine an effective way to measure spirituality without losing its very essence.

Given these concerns, staff developed explicit criteria with which to assess their potential evaluator candidates. Specifically, they looked for candidates who had one or more of the following characteristics:

- background/coursework with African-American researchers and practitioners who have argued for changes in how research and evaluation work is conducted in African-American communities;
- experience and a depth of understanding of African-American spirituality and the African-American church;
- a developmental and participatory approach that would encourage and support active participation of staff and participants in the evaluation design; and

- a strong methodological background, particularly around developing effective and innovative ways to measure intangible goals, such as increased spirituality.

As this example illustrates, many programs will have additional characteristics (besides those listed in this section) that they are looking for in an evaluator, given the specific type of program or participants served. Program staff and key stakeholders should feel free to explicitly determine these criteria in order to find the evaluator who is the best fit for the job. In some cases, it might mean hiring a team of evaluators, each with a different set of important characteristics. In other cases, it might make sense to couple an outside evaluator with an internal staff person, who has important perspectives about the clients being served, the community where the program operates, or the theories/assumptions upon which the program design is based. This is one of the reasons why we advocate that any program evaluator hired should work as part of a broader evaluation team made up of key staff and stakeholders. This will ensure that many perspectives are brought to bear on what is known about the program and what actions are taken based on this knowledge.

Example 2: In another example, an organization serving economically disadvantaged women hired an outside evaluator to evaluate the early stages of an innovative program for women living in poverty, being piloted in three communities. Not having a great deal of expertise on evaluation design at this time, the director and key staff hired the evaluators based primarily on their expertise in evaluation methods.

For several reasons, the relationship failed and the evaluation was not useful to project staff or other key stakeholders. The primary reason for the unsuccessful evaluation seemed rooted in the fact that project staff did not feel the evaluators hired understood how to work effectively with the disadvantaged women in the program. In addition, staff felt the evaluators had markedly different perspectives or values from staff about what constituted success or a positive outcome, and what was considered important to document. For example, the evaluators defined a positive outcome as securing employment where income taxes were withheld, while seeming to ignore outcomes the staff felt were critical, such as increased sense of self, self-esteem, coping and problem-solving skills, etc. At the same time, project staff did not feel empowered or knowledgeable enough about the technical aspects of evaluation to effectively make the case for why the evaluation would not yield useful information.

In addition, project staff did not understand their potential role in determining the purpose and goals of the evaluation, as well as defining the evaluators' roles. They assumed that this was the evaluators' job. Since the evaluators hired had also always assumed it was an evaluator's job to design the evaluation and define roles, roles and expectations were never explicitly discussed.

The evaluators hired had expertise in traditional survey research and impact studies, and so assumed their job was to judge the merit or worth of the program through traditional paper-pencil surveys and follow-up telephone interviews. Without an evaluation team where staff and other key stakeholders were empowered to contribute and shape the evaluation, critical information about the clients being served was never utilized. Ultimately, this impacted the credibility and usefulness of the evaluation results. For example, many of the women in the program lied on the survey in very obvious ways (e.g., noting they were not on welfare when, in fact, they were). Staff had been concerned that this would be a likely scenario on a paper-pencil survey, given the women's distrust of unknown forms they did not understand for people they did not know. Such experiences had almost always meant bad news in their lives.

Staff also knew that many of the women in the program did not have telephones in their homes and thus, follow-up phone interviews—an important second phase of data collection—were not going to be an effective means to collect information. However, because the staff and evaluators were not working together on an evaluation team, none of this information was utilized and acted on during the evaluation.

Finally, staff, who were very dedicated to the women being served and sensitive to the multiple barriers and difficulties they had faced, began to resent the evaluators for their seeming lack of sensitivity to the situations of these women, and how this evaluation might feel to them—cold and impersonal. The relationship quickly deteriorated, and when the report was finally completed, most staff members and participants were disappointed. They felt that the most profound aspects of the program were not addressed in the report, and that in many ways, the program and the women were unrecognizable. One staff member put it this way, "I didn't even see our program in this report. Or the women. Or really what the women got out of the program. None of that was in there."

In the second year of the program, with a chance to hire a new evaluator, staff members regrouped and decided the most important

characteristic to look for in an evaluator was someone who had values and a philosophy that matched that of the organization. They also wanted to hire an evaluator who was as much an advocate for the cause of women in poverty as they were. Finally, they wanted to ensure that this evaluator would treat the women served with the same high level of respect and care that each program staff person and volunteer did.

Two years later, working with an outside evaluation team with feminist and developmental evaluation principles, program staff felt much more positively about the role evaluation can play. They have also been proactive about determining the evaluators' roles as that of developmental consultants working with them to make continuous improvements, rather than independent judges of the merit or worth of their program.

Things to Remember . . .

- When hiring an external evaluator, consider his or her philosophical assumptions about evaluation and how appropriate they are to addressing the questions you want answered. In addition, invite finalists to meet project staff and others with whom they will be working closely to see who best fits with individual styles and your organizational culture.
- An important part of an evaluator's job (internal or external) is to assist in building the skills, knowledge, and abilities of other staff and stakeholders. It is better to have an evaluator who has spent time working with staff to integrate evaluation activities into day-to-day project management and delivery, than to have one who has conducted a perfectly constructed evaluation with strong recommendations that no one uses and with no one able to continue the work.
- Think of evaluation as everyone's responsibility. Be careful not to delegate all evaluation decision making to your evaluator. Stay involved and encourage teamwork.

Implementation Steps: Designing and Conducting an Evaluation

Implementation Steps

5. Determining Data-Collection Methods
6. Collecting Data
7. Analyzing and Interpreting Data

Evaluations must be carefully designed if they are to strengthen project activities. Evaluation designs that are too rigid, for example, can inhibit experimentation and risk taking, keeping staff from discovering more successful project activities and strategies. On the other hand, an evaluation design that is not carefully constructed can mask inherent biases and values; waste valuable resources gathering data that do not address important evaluation questions; or lead to inaccurate or invalid interpretations of data.

Following are three critical points to keep in mind throughout every phase of implementation:

Create a flexible and responsive design. The evaluation design should avoid procedures that require inhibiting controls. Rather, the design should include more naturalistic and responsive procedures that permit redirection and revision as appropriate.

You can maintain a flexible and responsive method of evaluation by:

- designing an evaluation that "fits" the needs of the target populations and other stakeholders;
- gathering data relevant to specific questions and project needs;
- revising evaluation questions and plans as project conditions change (e.g., budget becomes inadequate, staff members leave, it becomes obvious a question cannot be answered at this stage of the project);
- being sensitive to cultural issues in the community;
- knowing what resources are available for evaluation and requesting additional resources if necessary;

- understanding the existing capacity of the project (e.g., can project staff spend 30 minutes each day completing forms that document their activities and perceptions of the project?); and
- realizing the capabilities and limitations of existing technologies, and allowing time to deal with unforeseen problems.

Collect and analyze information from multiple perspectives. Our grantees grapple with complex social problems from a variety of cultural and social perspectives. Thus, the evaluation must be carefully designed to incorporate differing stakeholders' viewpoints, values, beliefs, needs, and interests. An agricultural project, for example, may produce noteworthy economic benefits, but have a negative impact on family and social culture. If you evaluated the impact of this project based only on an economic perspective, it would be deemed a success. However, is this a complete and accurate picture of the project?

Always return to your evaluation questions. Your evaluation questions (along with your ultimate purpose and goals) are critical to determining effective design. Too often, evaluation teams focus on the information and methods to collect information, and lose sight of the questions they are attempting to address. The more closely you link your evaluation design to your highest priority questions, the more likely you will effectively address your questions. Without this link, you risk collecting a great deal of information without shedding any light on the questions you want to answer.

Step 5: Determining Data-Collection Methods

The Foundation encourages the use of multiple evaluation methods, so projects should approach evaluation from a variety of perspectives. Just as no single treatment/program design can solve complex social problems, no single evaluation method can document and explain the complexity and richness of a project. *Evaluation designs should incorporate both qualitative and quantitative data-collection methods whenever possible.*

After deciding on the questions to address, you will need to decide what information is required to answer these questions, from whom and how the information can best be obtained. You will also need to decide how the information collected should be analyzed and used. Making these decisions early in the planning process will reduce the risk of collecting irrelevant information. While planning, remember to *keep your design simple, flexible, and responsive to the changing needs of the project*. Focus on the project-specific

appropriateness of data-collection methods, rather than on blind adherence to the evaluation design. By staying focused on the specific questions you want to address, you will make better decisions about what methods to use.

There are many different data-collection methods to choose from, including observation, interviews, written questionnaires, tests and assessments, and document review. (These are described in detail later in this section.) When deciding on which methods to use, consider the following four points:

1. Resources available for evaluation tasks:
Before devising an evaluation plan that requires a large portion of the project budget (more than 15 percent) and significant staff time, determine the resources available, and design your evaluation accordingly. By the same token, if you have budgeted less than 7 to 10 percent of your overall project budget for evaluation costs, you may consider requesting additional funds. Calculating the cost of several data-collection methods that address the same questions, and employing a good mix of methods, adequately thought out, can help stretch limited funds.

2. Sensitivity to the respondents/participants in the project:
Does the evaluation plan take into consideration different cultural perspectives of project participants and the evaluation team? For example, if half of the target population speaks only Spanish, do plans include printing client satisfaction surveys in both English and Spanish? Do you have evaluation staff fluent in Spanish to interpret the responses? Similarly, if your target population has a low level of educational attainment, it would be inappropriate to plan a large mailed survey that clients might find difficult to read and understand.

3. Credibility:
How credible will your evaluation be as a result of the methods that you have chosen? Would alternative methods be more credible and/or reliable, while still being cost effective? When deciding between various methods and instruments, ask the following questions:

- Is the instrument valid? In other words, does it measure what it claims to measure?
- How reliable is the measuring instrument? Will it provide the same answers even if it is administered at different times or in different places?
- Are the methods and instruments suitable for the population being studied and the problems being assessed?
- Can the methods and instruments detect salient issues, meaningful changes, and various outcomes of the project?

- What expertise is needed to carry out your evaluation plan? Is it available from your staff or consultants?

Increased credibility can also be accomplished by using more than one method, because the evaluator can then compare and confirm findings.

A final note on credibility: Earlier, we discussed the dominance of quantitative, impact evaluation designs. Many researchers, policymakers, funders, and other stakeholders may doubt the credibility of evaluations that are based on alternative paradigms (such as feminist, participatory, or constructivist models). For these types of evaluations, an important job of the evaluation team may be to educate stakeholders about the credibility of these evaluation designs and their effectiveness in addressing certain critical questions related to program success.

4. Importance of the information:
Don't forget to consider the importance of each piece of information you plan to collect, both to the overall evaluation and to the stakeholders. Some types of information are more difficult and costly to gather than others. By deciding what information is most useful, wasted energy and resources can be minimized.

Quantitative Versus Qualitative Methods
Most evaluations deal to some extent with **quantitative** information: things that can be counted and measured. For example, your evaluation may count the number of people involved in a project activity, the number of products or services provided, and the amount of material resources available or required. Other projects may measure the number of infant deaths in a community, the percentage of community board members who are minorities, the percentage of students who drop out of school, or the number of community residents living below the poverty line.

Qualitative information can be used to describe how your project functions and what it may mean to the people involved. A qualitative analysis may hold greater value than quantitative information because it provides a context for the project, and it may mean more to the project director who must make recommendations for improvement. Because qualitative information is full of people's feelings, it may give outside audiences a real understanding of the difference your project actually makes in the lives of people.

Project success depends, in part, on adequately considering hard-to-measure factors. For example, one may discover through quantitative methods that many households in a poor community are headed by single mothers. However, if one is unaware of the relationships among these households (e.g., cooperative child-

care arrangements, sharing resources, etc.), or the financial contributions made by male relatives, the importance of this information as an indicator of health status can easily be exaggerated. In other words, qualitative techniques, such as in-depth interviews and participant observation, can help the evaluator understand the context of a project. The context sets the framework for a meaningful understanding of other quantitative data (numbers, ratios, or percentages).

Determining the best way to collect data can be difficult. Following is a brief summary of principal methods for collecting data. Each describes the value and purpose of the method, types of information the method may collect, and includes an example of how the method could be used in combination with other methods. It is important to note that descriptions are not exhaustive, nor do we include summaries of all data-collection methods.

Observation

One way to collect information is to observe the activities of project staff and participants. Observation is especially useful when conducting context and implementation evaluation because it may indicate strengths and weaknesses in the operations of your project, and may enable you to offer suggestions for improvement.

Information gathered through observation will allow you to:

- formulate questions which can be posed in subsequent interviews;
- examine the project's physical and social setting, staff and clientele characteristics, group dynamics, and formal and informal activities;
- become aware of aspects of the project that may not be consciously recognized by participants or staff;
- learn about topics that program staff or participants are unwilling to discuss; and
- observe how project activities change and evolve over time.

Despite its value as a strategy for data collection, observation has limited usefulness in certain situations. For example, observing certain events, such as medical consultations, would be inappropriate if the observation violates the confidentiality of the doctor-patient relationship. Other types of observation, although legal, could violate cultural values or social norms. For example, a male evaluator of a project serving pregnant women should not intrude visibly on program activities if his presence would be disruptive. If in doubt, it is a good idea for the evaluator to talk with the project director about those situations he or she would like to observe.

An evaluator should also recognize that even the most passive, unobtrusive observer is likely to affect the events under observation. Just because you observe it, do not assume that you are witnessing an event in its "natural" state.

With these considerations in mind, here are a few tips:

1. Figure out how observation can be used to complement or corroborate the data you receive from other sources (e.g., surveys, interviews, focus groups). Establish goals for the observation, but be willing to modify them on site. You may even want to write some questions to guide your observations, such as: How do participants react to the project environment? Do they feel secure? Are staff members treated as equals? How do they address one another?
2. Be practical. If your time is limited and the project you are evaluating is large or scattered over several sites, you will not be able to observe everything. Decide what is most important or what cannot be learned from other data sources and concentrate on these areas.
3. Be systematic. Once you have established a focus (for example, staff relations), approach the matter from different angles. Observe the project at different times of day; observe a variety of different individuals; see if you can attend a staff meeting.
4. Be prepared. Develop instruments for recording your observations efficiently so that you can concentrate on what is happening on site. If the instruments do not work well, modify them.
5. Be inquisitive. Set aside some time to discuss your observations with the project director or other staff members. This will help you to put the events you observed in context and gain a better understanding of what happened.
6. Be open. *If you do not expect the unexpected, you will not find it.* —Heraclitus, 5th century B.C.

Interviewing

Interviewing, like other data-collection methods, can serve multiple purposes. It provides a means of cross-checking and complementing the information collected through observations. An evaluator interviews to learn how staff and clientele view their experiences in the program, or to investigate issues currently under discussion in a project. The inside knowledge gained from interviews can provide an in-depth understanding of hard-to-measure concepts such as community participation, empowerment, and cohesiveness.

Interviews can be used in all phases of the evaluation, but they are particularly useful when conducting implementation and context evaluation. Because interviews give you in-depth and detailed information, they can indicate whether a program was implemented as originally planned, and, if not, why and how the project has changed. This type of information helps policymakers and administrators understand how a program actually works. It is also useful information for individuals who may wish to replicate program services.

One of the first steps in interviewing is to find knowledgeable informants; that is, people who will be able to give you pertinent information. These people may be involved in service activities, hold special community positions which give them particular insights, or simply have expertise in the issues you are studying. One does not need a university degree or a prestigious title to be a valuable informant. Informants can be patients, staff members, community members, local leaders, politicians, or health professionals. Depending on the type of information you seek, you may interview one or many different informants.

In addition to finding informants, you must also decide which method of interviewing is most appropriate to your evaluation. Figure 6 describes different interviewing techniques, and highlights the strengths and weaknesses of each.

If you wish to record an interview, first obtain permission from the interviewee. If there are indications that the presence of the tape recorder makes the interviewee uncomfortable, consider taking handwritten notes instead. Tape-recording is required only if you need a complete transcript or exact quotes. If you choose to focus your attention on the interviewee and not take notes during all or part of the interview, write down your impressions as soon as possible after the interview.

Figure 6
Approaches to Interviewing

Type of Interview	Characteristics	Strengths	Weaknesses
Informal Conversational Interview	Questions emerge from the immediate context and are asked in the natural course of things; there are no predetermined questions, topics, or wordings.	Increases the salience and relevance of questions; interviews are built on and emerge from observations; the interview can be matched to individuals and circumstances.	Different information collected from different people with different questions. Less systematic and comprehensive if certain questions do not arise "naturally." Data organization and analysis can be quite difficult.
Interview Guide Approach	Topics and issues to be covered are specified in advance, in outline form; interviewer decides sequence and wording of questions in the course of the interview.	The outline increases the comprehensiveness of the data and makes data collection somewhat systematic for each respondent. Logical gaps in data can be anticipated and closed. Interviews remain fairly conversational and situational.	Important and salient topics may be inadvertently omitted. Interviewer flexibility in sequencing and wording questions can result in substantially different responses from different perspectives, thus reducing the comparability of responses.
Standardized Open-Ended Interview	The exact wording and sequence of questions are determined in advance. All interviewees are asked the same questions in the same order. Questions are worded in a completely open-ended format.	Respondents answer the same questions, thus increasing comparability of responses; data are complete for each person on the topics addressed in the interview. Reduces interviewer effects and bias when several interviewers are used. Permits evaluation users to see and review the instrumentation used in the evaluation. Facilitates organization and analysis of the data.	Little flexibility in relating the interview to particular individuals and circumstances; standardized wording of questions may constrain and limit naturalness and relevance of questions and answers.
Closed-Field Response Interview	Questions and response categories are determined in advance. Respondent chooses from among these fixed responses.	Data analysis is simple; responses can be directly compared and easily aggregated; many questions can be asked in a short time.	Respondents must fit their experiences and feelings into the researcher's categories; may be perceived as impersonal, irrelevant, and mechanistic. Can distort what respondents really mean or have experienced by so completely limiting their response choices.

From Patton (1990)

Group Interviews

An evaluator may choose to interview people individually or in groups. If the evaluator is concerned about maintaining the informants' anonymity or simply wants to make sure that they feel free to express unpopular ideas, it is best to interview people individually. This also allows the evaluator to compare various perspectives of an event, which is particularly useful when asking about sensitive topics. In Oscar Lewis's (1961) study of a Mexican family, *The Children of Sanchez*, each member of the Sanchez family was interviewed individually because he was afraid the children would not speak their minds if their domineering father was in the room.

When confidentiality is not a concern, and the evaluator is interested in quickly sampling the range of opinions on a topic, a group interview is preferable. One popular technique for conducting collective interviews is the **focus group**, where six to eight individuals meet for an hour or two to discuss a specific topic, such as local health concerns. Unlike a random sampling of the population, the participants in a focus group are generally selected because they share certain characteristics (e.g., they are diabetic, pregnant, or have other specific health concerns) which make their opinions particularly relevant to the study.

In the focus group session, participants are asked to respond to a series of predetermined questions. However, they are not expected or encouraged to work toward consensus or rethink their views, but simply to state what they believe. Many participants find the interaction stimulating and mention things they would not have thought of individually.

The **nominal group technique** and the **Delphi technique** are useful when you wish to pose a specific question to a group of individuals, and then ask the group to generate a list of responses and rank them in order of importance. These techniques were developed to facilitate efficient group decision making by busy executives, but they may also be useful in evaluation, particularly when groups composed of experts, community members, or project staff are making recommendations for ongoing projects.

Nominal group technique is a simple method that may be used if the individuals can be brought together at a face-to-face meeting. Delphi technique needs more elaborate preparation and organization but does not require the individuals to come together. Thus it may be useful if, for instance, you want community leaders to rank the most important local health problems but their schedules will not permit participation in a group meeting.

Examples of questions that could be answered with either technique include: What kinds of programs should be provided for adolescents? What are the most serious problems facing senior citizens in the community? Which program activity offered by this health project do you consider most successful?

Nominal Group Technique:
In this technique, five to nine participants (preferably not more than seven) sit around a table, together with a leader. If there are more participants, they are divided into small groups. A single session, which deals with a single question, usually takes about 60-90 minutes. The basic steps are:

1. Silent generation of ideas in writing—After making a welcoming statement, the leader reads aloud the question that the participants are to answer. Then each participant is given a worksheet (with the question printed at the top) and asked to take five minutes to write his or her ideas. Discussion is not permitted.

2. "Round-robin" feedback of ideas—The leader goes around the table and asks each member to contribute one of his or her ideas summarized in a few words. These ideas are numbered and written so they are visible to all members. The process goes on until no further ideas are forthcoming. Discussion is not permitted during this stage.

3. Serial discussion of ideas—Each of the ideas on the board is discussed in turn. The objective of this discussion is to obtain clarity and to air points of view, but not to resolve differences of opinion.

4. Preliminary vote—The participants are asked to select a specific number of "most important" items from the total list (usually five to nine). Then they are to rank these items on cards. The cards are collected and shuffled to maintain anonymity, and the votes are read out and recorded on a tally-chart that shows all the items and the rank numbers allocated to each.

5. Discussion of preliminary vote—A brief discussion of the voting pattern is now permitted. Members are told that the purpose of this discussion is additional clarification, and not to pressure others to change their votes.

6. Final vote—Step 4 is repeated.

Delphi Technique:
This technique is more elaborate than the nominal group technique. A series of mailed questionnaires is usually used, each one sent out after the results of the previous one have been analyzed. The process, therefore, usually takes weeks or months. The basic steps are:

1. Each participant lists issues in response to a question.
2. The evaluator develops and circulates a composite list of responses for members to rank in order of importance or agreement.
3. The evaluator summarizes results into a new questionnaire reflecting the group consensus.
4. Members again respond, rerank, and provide a brief rationale for their answers.
5. The evaluator generates a final questionnaire, giving revised priorities and rankings or ratings, and major reasons for dissent.
6. Each member makes a final evaluation for final group consensus.

Written Questionnaires
A survey or questionnaire is a written document that a group of people is asked to complete. Surveys can be short or long and can be administered in many settings. For example, a patient satisfaction survey with only a few questions could be constructed on a postcard and given to clients as they leave a clinic. More comprehensive responses may require a longer or more detailed survey.

Survey questions can be open- or closed-ended. Open-ended questions might ask: How do you feel about the program? What do you want to see happen in our community? Open-ended questions provide relatively rich information about a topic; allow participants to report thoughts, opinions and feelings; and are relatively low-cost. However, there are disadvantages. Sometimes people are reluctant to write down opinions, or the survey may be time-consuming to complete and analyze.

Unlike open-ended questions, closed-ended questions provide discrete, multiple-choice responses from which the respondent selects the most appropriate. For example:

How often do you use our center?
a. never
b. a few times a year

c. once a month
d. a few times a month
e. once a week
f. more than once a week

Closed-ended questions have the advantage of uniformity and easy translation for statistical analyses. Surveys can easily be administered to large groups of people and are usually easy to complete. However, closed-ended surveys tend to impose a set of fixed ideas or values on the respondent by forcing choices from a limited array of options. As a result, they are less likely to uncover surprising information, and they limit the emergence of in-depth understandings and nuances of meanings. Regardless of the type of questionnaire you use, written survey questions are inappropriate if the respondents have low language literacy or are unfamiliar with the conventions surrounding survey completion. A survey administered over the telephone or in person might be more appropriate for this population.

Here are a few simple rules to follow when developing a survey:

1. Make the questions short and clear, ideally no more than 20 words. Be sure to give the respondents all the information they will need to answer the questions.
2. Avoid questions that have more than one central idea or theme.
3. Keep questions relevant to the problem.
4. Do not use jargon. Your target population must be able to answer the questions you are asking. If they are not familiar with professional jargon, do not use it.
5. Avoid words which are not exact (e.g., generally, usually, average, typically, often, and rarely). If you do use these words, you may get information which is unreliable or not useful.
6. Avoid stating questions in the negative.
7. Avoid introducing bias. Slanted questions will produce slanted results.
8. Make sure the answer to one question relates smoothly to the next. For example, if necessary add "if yes...did you?" or "if no...did you?"
9. Give exact instructions to the respondent on how to record answers. For example, explain exactly where to write the answers: check a box, circle a number, etc.
10. Provide response alternatives. For example, include the response "other" for answers that don't fit elsewhere.

11. Make the questionnaire attractive. Plan its format carefully using sub-headings, spaces, etc. Make the survey look easy for a respondent to complete. An unusually long questionnaire may alarm respondents.
12. Decide beforehand how the answers will be recorded and analyzed.

After you have prepared your survey instrument, the next step is to pilot it. Ask your test audience to give you feedback on the clarity of the questions, the length of time needed to complete the survey, and any specific problems they encountered while completing the survey. Feedback from your pilot group will help you perfect the survey instrument.

Tests and Assessments

Tests and assessments can be useful tools in evaluation. In context evaluation, this method can allow you to gain information about the needs of the target population. In outcome evaluation, it can indicate changes in health status or behavior resulting from project activities. However, most of these measures require expertise and specialized training to properly design, administer, and analyze.

Physiological health status measures can be used to reveal priority health needs, or indicate the extent of particular health problems in a target population or community. Examples of these measures are broad-based screenings, such as cholesterol or blood-pressure readings, and physiological assessment data collected by other community organizations or hospitals. For example, a large number of low-birthweight babies reported by local hospitals may lead you to provide educational programs on prenatal care for prospective mothers.

Physiological assessments can also be used to measure the outcomes of a project. An increase in the birthweights of infants born to mothers in your prenatal care program is an indicator that the project may be answering an identified need. Statistical tests for significance can be applied to this kind of data to further confirm the positive effects of your project.

Knowledge or achievement tests can be used to measure participants' knowledge or behavior. Through testing before and after educational programs, you can assess what the participants need to learn, and then measure what they have actually learned. Be aware, however, that a person's knowledge does not prove that the person is using that knowledge in everyday life.

Another type of knowledge or achievement testing is done through observation, as when a staff member observes a mother interacting with her child to determine whether there has been progress in mastery of parenting skills. Or, a

home visit could include observation of improvement in an elderly person's mobility, or of specific skills learned by the caretaker. If they are to be useful as project outcome measures, these observations should be documented so that they can be compared across cases or across time. (See earlier section entitled Observation, page 73.) Standardized indices are available for coding observations.

Participant self-reports, including **standardized psychological and attitudinal assessments,** can also be used to measure need and assess outcomes. You may develop your own instruments to determine, for example, client satisfaction with existing health services or reactions to services offered by your project. Standardized questionnaires developed by health researchers on such topics as patient satisfaction, general health (including items on physical, emotional, and social function), mental health and depression, and disability status can also be used. There are advantages to using a questionnaire that has already been developed and field-tested. However, bear in mind that standardized assessments may not adequately reflect the important and unique aspects of your project or the situation of your target population.

Document Review

Internal documents are another source of potentially valuable data for the program evaluator. These include mission statements, organizational charts, annual reports, activity schedules, diaries, funding proposals, participant utilization records, promotional literature, etc. Such materials enable the evaluator to learn about the history, philosophy, goals, and outcomes of a particular project, and also provide clues about important shifts in program development or maturation. A document review may also be a good way to formulate questions for use in a survey or interview. Bear in mind that written documents do not necessarily provide comprehensive or correct answers to specific problems, as they may contain errors, omissions, or exaggerations. They are simply one form of evidence, and should be used carefully and in connection with other types of data.

Following is a list of some of the documents routinely kept by many projects, along with a few suggestions about how they might be used. The list is intended to be suggestive, not exhaustive:

1. Reports—These can be helpful in learning how the project originated, how it is currently organized, what it claims to do, how it intends to reach its objectives, the nature of its target population, what efforts are being made to achieve sustainability, etc.

2. Promotional literature—Brochures and flyers (or the absence of these items) can help the evaluator to assess the project's outreach efforts.

3. Logs and diaries—These may provide insights into project activities, staff relations, important events in the life of the organization, and changes in project development.
4. Minutes of meetings—These will generally provide information on attendance, dominance and other role relations, planning, and decision making.

Example: Staff of a large community-based organization dedicated to health reform initiated an evaluation designed to determine the impact of, and lessons learned from, their past and present programming. Their evaluation did not focus on the "success" of individual projects, but rather concentrated on the organization's progress as a whole in implementing its overall mission of improving community residents' access to health care, increasing their knowledge of prevention-focused health care, and reducing the occurrence of high-risk health behaviors.

An evaluation team, consisting of representatives of each project and an external evaluator, previously determined the primary evaluation questions: Have community residents experienced greater access to health-care services overall? Do community members have increased knowledge as to what constitutes high-risk health behaviors? Have high-risk health behaviors decreased among community members? What lessons have been learned from implementing these programs? What opportunities have been missed in improving the overall health of community members?

Given the purpose, key questions, and human and financial realities of what was feasible, the evaluation team determined that based on the significant amount of data which individual projects had already collected, data-collection methods selected for this evaluation needed to maximize the use of existing data so as not to duplicate efforts and waste precious resources. In order to obtain information from a variety of perspectives, however, the evaluation team decided that the questions they wanted to answer would be best addressed through a combination of quantitative and qualitative data obtained from on-site observations and interviews, as well as a review of previously collected data.

On the qualitative side, they planned an extensive review of existing documents, such as project-specific interim reports, previous evaluations, mission statements, and organizational charts. They anticipated that this review would provide the evaluation team with the context of each

project's history, goals, and achieved outcomes in relation to the organization as a whole. In addition, based on this information, they intended to identify key informants for subsequent interviewing and on-site observation purposes.

On the quantitative side, since each project was required to collect data on an ongoing basis in relation to number of participants/clients served, as well as various project-specific information data (such as length of time participants spent in a project, number of referrals given, pre- and post-program impacts, etc.), some data were already available for purposes of this evaluation and did not need to be collected again.

By employing these data-collection methods, the evaluation team intended to obtain a more complete picture of this organization's health reform efforts to date.

Things to Remember . . .

- Determine data-collection methods based on how appropriate they are for answering your key evaluation questions and for achieving the ultimate purpose of the evaluation.
- Tie method selection to available resources. This may mean revising your evaluation design and methods, or determining other options to stay within budget. It may also mean finding additional resources to fund the evaluation design which will be most effective and useful.
- Choose methods based on what is appropriate for the target population and project participants.
- Strengthen the credibility and usefulness of evaluation results by mixing evaluation methods where appropriate.

Step 6: Collecting Data

Once you have refined your evaluation questions and determined what evaluation methods to use, you and your evaluation team are ready to collect data. Before spinning your wheels developing interview guides and survey questionnaires, examine the existing information about your target population, community, or project. Important questions to ask of your organization and other sources include:

- Why do you collect this information?
- How is it currently used?
- Can it help you address your evaluation questions? How?
- What is still missing?
- Are there other sources of information for what is missing?

Remember to "collect only the information you are going to use, and use all the information you collect."

Since the Foundation supports an integrated, rather than a stand-alone approach to evaluation, the data-collection step becomes critical in the project-level evaluation process. Most organizations collect a great deal of information on a range of topics and issues pertaining to their work. However, that information often stays unused in management computer systems, or remains isolated in a particular area of the organization. Therefore, we see part of the data-collection process as examining existing tracking systems, deciding why certain data are collected and how they are used, and thinking critically about what kinds of data you and staff need but have not been collecting in consistent ways. Discussions about in-house data collection also encourage communication across functional lines so that relevant data, collected at different points in the system or process, can be connected, leading to new insights, actions, and positive changes.

As with all of the steps in this evaluation blueprint, it is important to connect this phase with the others in the process. Many evaluators get caught in the information-collection step—collecting everything they can get their hands on and creating more and more instruments with which to collect it. Social programs and services are becoming increasingly complex to handle increasingly complex situations. With so many types of data and information, it becomes difficult to stay focused on the specific questions that you are trying to address. Encourage those responsible for data collection to continually ask themselves how each piece of data they collect will be used, how it will fit with the other pieces of data, and how it will help answer the questions at hand.

The data-collection step also helps you and your evaluation team to revise the design and methods based on resources (financial and human); to examine how the evaluation process is received by clients and other people from whom information is collected; and to assess the usefulness of the information collected.

> ***Example:*** In the previous example describing a community-based organization dedicated to health reform, the evaluation team, composed

of both project staff and evaluators, decided to start by collecting existing data for several reasons. First, given the scope of the evaluation, the team did not have access to the human and financial resources necessary to obtain such a vast amount of data from scratch. In addition, since this organization is large and has a considerable number of projects, data presented in a variety of formats were readily available. Finally, the evaluation team believed that starting the data-collection process with the examination of existing data would provide important information needed to guide the second phase of data collection—interviews and on-site observations.

The process of collecting existing data was helpful in terms of refining evaluation questions, identifying key informants for subsequent interviewing purposes, developing interview protocols, and determining what data important to the evaluation were missing. As a result, the evaluation team was better able to tailor the interviewing and on-site observation phase of data collection to fill in gaps in the existing data. For example, prior to this review, volunteers had not been identified as a group to be interviewed. After having reviewed existing material, however, the evaluation team became aware of the crucial role volunteers played in most projects. They were subsequently included in the interview schedule and provided vital information in terms of the evaluation.

An explicit goal of the second phase of the data-collection process was to speak with a cross section of people who represented many layers of a project. Given the limited time, money, and available staff for interviewing purposes, however, the evaluation team was unable to interview all of the people identified in phase one. With this in mind, the evaluation team determined that it was especially important to interview community members and direct line staff for purposes of confirmation and clarity. Having done that, the evaluation team felt that they got a much fuller picture of what occurred, what worked well, and what the challenges of each project—as well as the organization as a whole—were, despite interviewing fewer people than originally planned.

By working as a team, members of the evaluation team were able to prioritize together what was doable in terms of collecting data, given existing constraints. Also as a result of this teamwork, multiple perspectives were contributed to the data-collection process, and a greater sense of ownership of the evaluation as a whole was evident throughout the organization.

This two-phased data-collection process provided important lessons for improving the organization's existing data tracking systems. It provided staff with the opportunity to identify what kind of data they needed but were not collecting in an ongoing, systematic way, as well as any irrelevant data they were inadvertently collecting.

In addition, the process of interviewing key stakeholders in the projects not only helped address key questions specific to this evaluation, but also provided additional information such as staff and volunteer development needs, as well as program improvement ideas. Based on this experience, the evaluation team decided to integrate periodic interviews into their ongoing program tracking and data-collection system. This has enabled them to simultaneously provide staff development opportunities, work toward program goals, and collect data on an ongoing basis.

Things to Remember . . .

- Collect only the data you will use and that are relevant to your evaluation questions and purposes.
- Involve all staff involved in the data-collection phase in up-front question formation.
- Revise data-collection strategies based on initial analyses. What is working? What is not working? What pieces of data are still missing?
- Base changes to existing tracking/data-collection strategies on what is learned from evaluation.

Step 7: Analyzing and Interpreting Data

After designing an evaluation and collecting data, the information must be described, analyzed, interpreted, and a judgment made about the meaning of the findings in the context of the project. This process can be complicated and, at times, technical. In fact, many books are dedicated to the many methods of evaluation. Thus, it is not possible for an introductory manual to adequately explain the techniques of analysis and interpretation. In the following pages, however, we summarize some of the basic techniques for organizing and analyzing data.

Quantitative Analysis

Most often we think of statistical or quantitative analysis when we think about analyzing. Project staff without a background in statistics may be intimidated by quantitative analysis; what we often see is that deference is given to external or internal evaluators because they know how to "do evaluation." However, there are ways that project staff and an evaluation team without strong statistical backgrounds can analyze collected data. For example, you can begin by converting quantitative findings (e.g., numbers from utilization records, or answers on questionnaires) into percentages or averages.

The importance of valuing and seeking multiple perspectives comes into play during this phase of the evaluation. Quantitative data analysis does require interpreting the results and seeing if it makes sense given the project's contextual factors—factors that staff know better than most. Project staff and the evaluation team should work together and ask: Do these results make sense? What are some possible explanations for findings that are surprising? What decisions were made about categories and indicators of success? Have we missed other indicators? How might what we chose to collect and analyze be distorting the program/initiative? And most importantly, how will the numbers and results help us decide what actions will improve the program?

Remember that we want evaluation to support programs and help them improve. *Complex statistical analyses of a well-designed experimental investigation that does not lead to improvements are less desirable than a thorough but simple statistical analysis of existing tracking records that leads to positive changes in both the program and in the tracking system.*

Qualitative Data Analysis

Qualitative data includes information gathered from interviews, observations, written documents or journals, even open-ended survey questions. Information gathered from interviews and observations is often recorded in lengthy narratives or field notes. In some cases, interviews are tape-recorded and then transcribed. Some of these accounts are useful and can stand alone—providing important information about how the program is working. In most cases, however, it is valuable to analyze your qualitative data in more systematic ways.

The Foundation, advocating for qualitative methods and analysis as a way to better understand programs, feels that not enough people understand the power and logic of qualitative methods. This is, in large part, because they have not been trained to systematically analyze qualitative data. Too often, qualitative data are

seen as nice anecdotal information that bring the real results (the numbers) to life and put them in context. However, qualitative data help explain how a program works and why it has played out in a certain way, why a program faced certain stumbling blocks, and may even explain—and provide evidence of—those hard-to-measure outcomes that cannot be defined quantitatively.

As previously noted, there are many subtle nuances to qualitative data analysis that cannot be discussed in this manual, but there are many resources available for those interested in using qualitative analysis to strengthen their evaluation findings. The following describes some basic qualitative analysis techniques.

Categorization and Coding Techniques:
Qualitative data allows you to look for similarities across several accounts, interviews, and/or documents. Examining interview transcripts, observation field notes, or open-ended surveys for patterns and themes involves categorizing your notes into recurring topics that seem relevant to your evaluation questions. This is often done by first reading through your materials to identify themes and patterns. The next step is to cut up transcript copies, sorting by the key categories you discovered, and affixing these pieces to index cards (always remembering to identify where they came from) or to use computer programs that perform the task electronically. Organizing your material in such a way will make it easier to locate patterns, develop new hypotheses, or test hypotheses derived from other sources.

Contextualization Analysis Techniques:
Although using categorization techniques is a powerful way to document patterns and themes in a program, unless used with contextualization techniques which focus more on how things fit together, categorizing can lead to premature generalizations about the program. Case studies and narrative summaries about a particular piece of the program or participant are contextualization techniques that preserve and clarify the connections. These techniques bring to light important contextual factors and individual differences which are often hidden from view when we break transcripts and qualitative data into disconnected categories.

Memo-Writing Techniques:
Many critics of qualitative evaluation methods argue that it is too subjective; that evaluators lose their objectivity when they get close to a program or people involved. Others would argue, and we at the Kellogg Foundation would agree, that complete objectivity is not possible even in quantitative work. Furthermore, we believe that professional qualitative evaluators have devised effective ways to deal with subjectivity—by reflecting on their own values and biases, and then

analyzing how these affect what information they collect and don't collect, what they hear and don't hear, how they interpret the data, and what conclusions they ultimately make. Through an ongoing process of writing reflective memos about the evaluation process, their data, and their interpretations, qualitative evaluators ensure that they pay attention to the influences of biases and values, an inevitable part of any evaluation. This further supports our suggestion that evaluations be planned and conducted by an evaluation team. In this way, memos can be shared, and multiple perspectives come together to ensure that all perspectives have been considered.

Other Forms of Analysis

Finally, we want you to know that there are additional analysis techniques which are based on other evaluation philosophies. For instance, feminist research methods have led to analysis techniques such as Carol Gilligan's voice-centered analysis (Brown and Gilligan, 1990; Brown and Gilligan, 1992; Gilligan, Brown, and Rogers, 1990). Demonstrating that traditional interview analysis techniques missed or misrepresented critical stories that girls and women were telling about their moral development, Gilligan developed her theory of analysis to bring some of these hidden stories and themes to the surface.

Participatory evaluators have developed techniques where the participants work with evaluators to analyze their own interview transcripts, and together, develop interpretations.

The point is that many forms of analysis besides statistical analysis exist to help us understand and explain what is happening with social programs and services today. Which analysis techniques or combination of techniques to use depends again on the particulars of your project, who you are serving and in what contexts, and the questions you are attempting to answer.

Time pressures and constraints associated with conducting evaluations often limit an evaluator's ability to conduct thoughtful and in-depth analyses. We believe it is important to invest enough time and resources in the analysis and interpretation step, since it is during this integral phase that decisions are made and actions taken.

In summary, interpretation involves looking beyond the mounds of raw data to ask important questions about what the results mean, what led to the findings, and whether the findings are significant. Remember to involve stakeholders as your evaluation team seeks answers to these questions. Besides reducing their anxiety, you will gain insight from their knowledge about the program and maintain excitement about the evaluation process.

Example 1: Staff of a program focused on increasing public education and awareness related to groundwater quality and drinking water issues hit the ground running with a series of workshops aimed at providing local government representatives and community residents with information and knowledge about relevant groundwater issues. Midway through the workshop series, staff realized that they were not doing what they set out to do. In particular, they were not reaching their intended audience—local government officials and community residents—and instead found themselves "preaching to the choir"—other service providers and advocate organizations like themselves.

Wanting to understand why local officials and average citizens were not coming to the workshops as they had marketed them, the project director hired a local evaluator to work with two staff people to learn more. The team constructed customer surveys and then interviewed a sample of local government officials as well as community residents to determine how, more effectively, to structure and deliver the workshop series.

This evaluation team conducted the interviews and then worked together on the analysis and interpretation phase. Through a series of meetings, the team dialogued and analyzed the results of the customer interviews. As a team, they were able to check out their assumptions, act as sounding boards for particular interpretations, and build consensus based on their different perspectives. In addition, the staff members on the evaluation team who were responsible for the curriculum development and implementation were able to connect the customer interview results to implications for curriculum revisions. Because they were actively involved in the analysis and interpretation phases of the evaluation, the two staff members had a real sense of ownership over the changes they decided to make to the curriculum and understood, firsthand, how these changes connected to the data collected from their customers.

Ultimately, as a team, they were able to derive many insights about the strengths and weaknesses of the workshop series. They learned what motivated their customers and why they initially came to the workshops. Based on their collective interpretation of the customer interviews, staff members (with the evaluator's input) were able to redesign the workshop series to meet the customers' needs and interests.

Another overarching insight of the evaluation was the fact that they had defined their geographic area too largely. People were motivated by issues affecting their immediate community and neighborhoods. The

evaluation team found they needed to frame the issues in terms of the people's lives and landmarks. Workshops originally designed for almost a dozen local communities were split into smaller areas of only a few local communities that were related to one another, and focused on topics directly affecting those communities.

The analysis of the customer interviews also yielded important information about how to structure the workshops; what audience mixes made the most sense; the timing and locations which would be most effective; and the topics which were most relevant to particular communities.

For each lesson learned from the analysis and interpretation of the surveys, the staff members made relevant changes to the content and format of the workshops. Through additional analysis of post-workshop evaluation surveys and informal follow-up calls, they found these changes yielded both increased attendance, and a stronger educational vehicle for the goals of the program. This next phase of analysis and interpretation of post-workshop data is also providing the evaluation team with important new questions to address, including: Should this workshop be an ongoing series over a longer time period? Are participants interested in creating a network and problem-solving group out of this workshop experience around groundwater issues? How might we maintain and support that network?

Example 2: In the case of the national organization serving disadvantaged women, there is a good example of how a change in evaluation questions and primary data collection methods led to new interpretations and findings. These new interpretations, in turn, led to important changes in organizational priorities.

As we discussed in an earlier section, the first evaluation focused on whether a pilot program for disadvantaged women worked. Designed as a traditional impacts study, the evaluation consisted primarily of client surveys and structured follow-up interviews—effective and appropriate methods to address the primary evaluation question: Did the program work?

However, staff realized after reading the evaluation report that although they did want evidence of whether the program was successful, there were many other critical questions which were not addressed in this first evaluation process. When a second evaluation was commissioned for the following year, staff decided it was important to work with the evaluators

to determine what they wanted to know about their program and organization. Then, the evaluation could be designed specifically around these questions. As a team, with the evaluators' input and facilitation, staff decided that this pilot program marked an important shift in their organization's identity and how they went about achieving their mission. They wanted the evaluation to address the following questions:

- How and why did this program work?
- Who were the women they served, really?
- What did this program mean to the women participating?
- How did this program connect to the organization as a whole?

In addition, they wanted to ensure that the evaluation would provide them with useful information to improve their programs, services, and organizational structures.

This process provides a good example of how a change in evaluation questions can lead to the selection of different evaluation methods, and ultimately to different interpretations. Given the revised evaluation questions, the evaluation team determined that a primarily qualitative approach was the most appropriate, at least for the first year. In order to answer their questions effectively, the evaluation team needed to understand 1) the meaning of related events and experiences to those involved in the program; and 2) the contextual and situational factors that impact human and organizational development change. This is best addressed through qualitative methods, including in-depth interviews, focus groups, and observation.

Once the data were collected, the evaluation team began the analysis and interpretation of each particular component (e.g., interviews, observations, document review) with close reading of the transcripts and field notes to look for themes and connections, and determine the meaning. Then they analyzed the data from interviews, focus groups, and observations using traditional categorization and coding techniques, looking for patterns and themes both within and across the program sites.

At the same time, the differences across sites and among the different perspectives helped the evaluation team to avoid generalizing too quickly about patterns, themes, and hypotheses. To ensure that they did not lose site of individual differences and contextual factors, the evaluation team combined categorizing analysis techniques (which focus on similarities)

with contextualization techniques, such as narrative summaries and mini case studies. These techniques helped them preserve the connections within a particular site or person's experience and bring to light contextual factors and individual differences that are often hidden from view when using only categorization techniques.

Ultimately, these changes in data-collection methods and analysis techniques led to new interpretations about what was going on. Rather than learning primarily about the program and whether it had "worked," the second evaluation led to new and important insights about the transformations this organization was going through as a whole. The evaluation provided information about the pilot program in the context of larger organizational changes. The evaluation team learned that the pilot program marked an important transition in their organizational identity and in how they delivered services to their client base. It also was an important time of testing out and solidifying their organizational principles and values based on what they had learned about what worked for women in poverty. They also learned that staff and volunteers at all levels of the organization were feeling stress and confusion as a result of the recent changes.

These new interpretations, in turn, led to important strategic actions, including an increased investment in staff development to support staff through this change process and a focus on articulating the organization's values and principles and a logic model/theory of change about how the organization intended to achieve its mission to reduce the number of women in poverty.

Things to Remember . . .

While analyzing and interpreting both quantitative and qualitative data, be careful to avoid the following pitfalls:

- Assuming that the program is the only cause of positive changes documented. Several factors, some of which are unrelated to project activities, may be responsible for changes in participants or in a community. It is usually not possible to isolate impacts, and the evaluation report should at least acknowledge other factors which may have contributed to change.
- Forgetting that the same evaluation method may give different results when used by different people, or that respondents may tell the evaluator what they believe he or she wants to hear. For example, two interviewers may ask the same questions but receive different answers because one was friendlier or more patient than the other. Real problems or difficulties may be ignored or hidden because people want the project to succeed or appear to be succeeding.
- Choosing the wrong groups to compare or comparing groups that are different in too many ways. For example, gender, age, race, economic status, and many other factors can all have an impact on project outcomes. If comparisons between groups are important, try to compare those with similar characteristics except for the variable you are studying.
- Claiming that the results of a small-scale evaluation also apply to a wide group or geographic area. For example, it is misleading to evaluate participants' responses to a particular intervention in one city and then claim that the results apply to the U.S. as a whole. While this may well be the case, an evaluation report should reflect only the data analyzed.

Utilization Steps: Communicating Findings and Utilizing Results

Utilization Steps

8. Communicating Findings and Insights
9. Utilizing the Process and Results of Evaluation

Step 8: Communicating Findings and Insights

Relevant stakeholders and your evaluation team should discuss the most useful ways to communicate evaluation findings. This communication might take the form of weekly discussions with the evaluation team. It might include monthly discussions or roundtables with a larger audience. Project staff might ask external evaluators who are not on site to send biweekly memos on their insights and reflections for response and comment. The critical point is to involve everyone who will need this information in discussions about how best to communicate the progress of the evaluation.

A commitment to ongoing dialogue and more interactive forms of communication will not only increase ownership and motivation to act on what is learned, but will also assist in refining the evaluation design, questions, methods, and interpretations. This iterative process will optimize resources to, in turn, answer the most pressing questions.

Marketing and Dissemination

We strongly suggest that marketing and dissemination planning be integrated with evaluation planning. Ideally, the information that will inform the marketing and dissemination function should be considered early on so that the data collection plan focuses on obtaining relevant data. In addition to the ongoing communication mechanisms you set up during the evaluation process, it is important to develop more formal reports and presentations that provide information about your project, including evaluation findings. These more formal reports should be disseminated in a number of ways, and to a variety of audiences.

Writing Annual and Final Reports to the Kellogg Foundation:
Annual and final reports to the Kellogg Foundation are the primary ways we learn from projects' experiences. They are most helpful in informing our future grantmaking when they document what is happening or has happened in your project and why, and describe the lessons learned in the process of designing and implementing your project. Yet, many such reports simply document project activities and expenditures, and say little about lessons learned and significant outcomes.

The best initial source to consult in preparing your reports to the Kellogg Foundation is the commitment letter which informed you of our decision to provide a grant to your organization. As mentioned earlier, this letter typically lists the significant evaluation questions that we believe your project may help answer. While neither your evaluation nor your reports should be limited solely to these questions, they do provide a good starting point.

Annual and final reports need not be lengthy. In fact, a concise, well-written report of ten pages is more likely to influence our programming than one hundred pages of raw data.

Communicating to Other Stakeholders:
Disseminating information about your project to outside audiences can serve many purposes: improve the functioning of related projects and organizations; provide an accounting to funding and regulatory bodies; convince diverse audiences of the project's importance; and generate further support for the projects you have implemented. Evaluation findings presented in the media, describing project activities and the conditions of the target group, can increase local understanding and involvement.

Be creative and innovative in reporting evaluation findings. Use a variety of techniques such as visual displays, oral presentations, summary statements, interim reports, and informal conversations. Additional ideas include:

- Write and disseminate a complete evaluation report, including an executive summary and appropriate technical appendices.
- Write separate executive summaries and popular articles using evaluation findings, targeted at specific audiences or stakeholder groups.
- Write a carefully worded press release and have a prestigious office or public figure deliver it to the media.
- Hold a press conference in conjunction with the press release.

- Make verbal presentations to select groups. Include demonstration exercises that actively involve participants in analysis and interpretations.
- Construct professionally designed graphics, charts, and displays for use in reporting sessions.
- Make a short video or audiotape presenting the results, for use in analysis sessions and discussions.
- Stage a debate or advocate-adversary analysis of the findings in which opposing points of view can be fully aired.

Things to Remember . . .

Most evaluations will involve the preparation of a formal written report. When writing reports, keep the following in mind:

- Know who your audience is and what information they need. Different audiences need different information, even when addressing the same issues.
- Relate evaluation information to decisions. Reports written for decision-making purposes should first state the recommendation, followed by a summary of the relevant evaluation findings.
- Start with the most important information. While writing, imagine that your audience will not have time to read the whole report; be brief, yet informative. Develop concise reports by writing a clear abstract and starting each chapter, subsection, or paragraph with the most important point.
- Highlight important points with boxes, different type sizes, and bold or italic type.
- Make your report readable. Do not use professional jargon or vocabulary that may be difficult to understand. Use active verbs to shorten sentences and increase their impact. Write short paragraphs, each covering only a single idea.
- Edit your report, looking for unnecessary words and phrases. It is better to have someone else edit your work; if you must edit yourself, allow a day or two to pass between writing and editing.

Step 9: Utilizing the Process and Results of Evaluation

Above all, an evaluation must provide usable information. It must enable project directors, for example, to guide and shape their projects toward the greatest effectiveness.

The final action step we want to discuss is how we use the process and results of evaluation. Here is where so many evaluations fall short. Many grantees complain that evaluations of their programs are not used to make decisions; oftentimes, they do not provide information that is useful in the day-to-day management of the program. In other cases, although staff have found evaluations useful, it is hard to determine exactly when and how the evaluation process or results led to decision making or actions. Evaluation often gets things started but only in incremental steps.

We believe that one of the most important characteristics of an effective evaluation is that it does provide usable information—information that project staff and other stakeholders can utilize directly to make decisions about the program. An evaluation report that sits on someone's shelf will not lead us to improved program design and management. Effective program evaluation supports action. Useful evaluation processes and results inform decisions, clarify options, identify strengths and weaknesses, and provide information on program improvements, policies, and key contextual factors affecting the program.

It is important that the evaluation team and other stakeholders think about the question of use from the outset of the evaluation process. If you wait until the end of the evaluation to discuss how you want to use evaluation results, it will be too late; the potential uses of the study will already have been determined by the decisions made along the way. Therefore, during each planning and implementation step—Identifying Stakeholders, Developing Questions, Budgeting, Selecting an Evaluator, Determining Methods, Collecting, Analyzing, and Interpreting Data, and Communicating Findings—you should engage in discussions about how you want to use the evaluation process and results to make decisions and take actions.

Start these discussions about use by asking questions such as:

- What do you, other staff, and key stakeholders need to know more about?
- What decisions do you feel you need to make, but need more information?
- What will you do with the answers to your questions? (Play out different scenarios, depending on the different answers that you may find.)

Then articulate (in writing) how you and your evaluation team intend to utilize evaluation results. Be as specific as you can and revise as you go through the process. By determining your priority uses early in the process, you will be able to make more effective decisions about design and methodology questions, and will end up with information you need to make the decisions you planned to make.

Additional questions to discuss throughout the process include:

- Who will make the decisions and when?
- What are the different issues that are likely to surface related to these decisions?
- How are decisions made in this organization?
- What other factors may affect the decision-making process?
- How will we know if we used the evaluation results and process as we planned?

Staff and stakeholders are more likely to use evaluation if they understand and have ownership over the evaluation process. Therefore, the more people who have information about the evaluation and have been actively involved in the process, the easier it will be to facilitate using the process and results for program improvement and decision making. This is why we recommend forming an evaluation team and developing mechanisms for spreading the word about the evaluation process to other staff and stakeholders.

However, it is not enough to engage in discussions about how you will use evaluation and to ensure that everyone understands the benefits of evaluation and is actively involved in the process. Many programs have taken these initial steps and were still unable to use evaluation to support action. One of the reasons for this is that there are many individual and organizational obstacles to using information and testing assumptions about a program's effectiveness. Obstacles include: the fear of being judged, concern about the time and effort involved, resistance to change, dysfunctional communication and information-sharing systems, unempowered staff. Address this issue by engaging in discussion and reflecting about the specific obstacles to using information within your organization.

Using Evaluation Findings

Specific uses will depend on the overall purpose of your evaluation and the questions you are attempting to address. The following highlights several specific uses of evaluation findings:

Improving Your Program:
A goal of every evaluation should be to improve the program, and evaluation findings should support decisions and actions about how best to do so. Specific findings might be used to identify strengths and weaknesses of your program or provide strategies for continuous improvement. You may decide to focus evaluation questions on organizational issues. In this case, findings could lead to strategies that would help staff manage more effectively, improve organizational culture or systems, or improve staff interactions and relationships with clients.

Evaluating the Effectiveness of a Program:
In cases where the primary purpose of evaluation is to assess the effectiveness of the program, evaluation findings should support decisions around accountability and quality control. In some cases, findings might be used to decide a program's future, determine the likelihood of continued funding, or make decisions about program expansion.

Generating New Knowledge:
Another potential goal of evaluation is to discover new knowledge about effective practice. In particular, the Kellogg Foundation advocates focusing evaluation questions on how and why programs work, for whom, and in what circumstances. Findings and insights from addressing these types of evaluation questions provide important information about general principles of effective practice, cross-cutting themes, connections between underlying theories and practice, and sometimes lead to new and enhanced theories about human and organizational development. In addition, these types of findings can be used to collaborate, share, and learn across programs and initiatives with common themes and principles. They are often at the heart of policymaking decisions, as well.

Utilizing the Evaluation Process
We believe that grantees can learn a great deal from the evaluation process, as well as from evaluation results. This is particularly true when project staff explicitly discuss what they want to learn from the evaluation process at the beginning of the evaluation.

An evaluation process provides multiple avenues to impact staff, volunteers, participants, and other stakeholders in positive ways. Those involved in the evaluation will learn how to recognize the thinking processes and values that underlie their particular approach to evaluation. They will gain skills in building consensus; identifying key questions; collecting, analyzing, and interpreting data; and communicating results and insights. The process will help staff become more focused and reflective. An effective evaluation process can lead to positive changes

in organizational culture and systems. It can also increase program participants' sense of self-efficacy and self-esteem or assist in challenging misconceptions about a particular target population. The following highlights several specific uses of the evaluation process:

Building Shared Meaning and Understanding:
One way you can use the evaluation process is to improve communication and shared understanding between different groups involved in the program (e.g., between line staff and managers or between staff and volunteers). Some evaluations can help key stakeholders, or the community in general, better understand the target population being served, particularly disenfranchised groups who are not often heard. In short, the evaluation process can provide a way to better connect all those involved in the program/initiative, and to build on these collective understandings.

Supporting and Enhancing the Program:
When data collection and analysis are integrated into program design and implementation, the evaluation process can actually become part of the program intervention. For instance, one program serving women in prison uses a Life Map where participants write the history of their life and then present their Life Map to the class. This Life Map is used to collect important baseline data about each woman's history and life situation; however, it is also a critical part of the ten-week intervention. It becomes a time for women to reflect on their lives and how current situations might have been affected by past events and circumstances. It is also an important trust-building activity as participants begin to share and open up with one another. So, while this activity provides valuable data for evaluative purposes, it also strengthens the program's impact.

Traditional evaluation proponents would consider this type of evaluation process and use highly problematic. From their perspective, evaluators are supposed to remain objective third parties who do not engage in or influence program implementation. However, it is in line with our view that evaluation should be an integral, ongoing part of the program, rather than a separate, stand-alone piece. In addition, this type of evaluation process helps increase the likelihood that evaluation findings are used, since staff are actively involved in the data-collection process.

Supporting Human and Organizational Development:
No matter what your role in a program, being involved in an effective evaluation can impact your thinking and interactions in positive ways. Staff, volunteers, participants, and other stakeholders involved in the evaluation will have opportunities to acquire important skills from the process, including: identifying

problems; reflecting; setting criteria; collecting, analyzing, and interpreting data; debating and determining alternative solutions.

With your evaluation team, think explicitly about how you can help increase the impact that participatory evaluation processes can have on staff, volunteers, and especially program participants. At its best, these types of evaluation processes become capacity-building processes, where different groups involved in the program discover and build on their assets and skills.

> ***Example 1:*** An initial evaluation of a program designed to provide educational services to families and children in an economically disadvantaged urban community helped project staff discover that they were operating the program based on a set of implicit and unspoken assumptions. The fact that these assumptions were not put in writing or discussed explicitly as part of the program seemed to be contributing to problems with new staff's ability to understand the program, its goals, and underlying principles. Founding staff used these evaluation insights to create a historical overview outlining the origins of the project, the path it had taken to get where it is, and the assumptions underlying the program and its mission. This historical overview is now used during the orientation of new staff, as well as in presentations of the program to potential partners and key stakeholders, to help others understand the history and key principles of the program.
>
> To continue to use evaluation information effectively to improve the program, staff are currently using the historical overview as the basis for creating a program logic model, which will provide further details about how the program works to achieve its goals. The process of developing a program logic model has provided further opportunities for staff and key stakeholders to build consensus and shared understandings about the program and how it works. In addition, through the process of collecting data around the different pieces of the logic model, staff will be able to determine strengths and gaps in the program and make quality improvements. It is through this iterative process of evaluation inquiry followed by the use of evaluation findings that grantees will best be able to learn about and continuously improve their programs.
>
> ***Example 2:*** In a previous example where an organization had one negative and then one positive evaluation process, a critical factor in making the second evaluation process successful was the high priority placed on supporting human and organizational development. Defined as a developmental evaluation focused primarily on supporting human and

organizational development, the evaluation was designed to be interactive, collaborative, and to build on the skills and assets of staff, volunteers, and participants.

Staff were re-introduced to the concept of evaluation and encouraged to share any negative feelings they had developed based on their first experience with evaluation. An evaluation team consisting of key staff members, volunteers, and a board representative actively participated in defining evaluation questions, determining methods for collecting data, and discussing the meaning of the data and potential actions to be taken. Staff learned about evaluation processes and techniques and began to take ownership over the evaluation process in ways they had not in the first evaluation.

In addition, interviews and focus groups with program participants were designed as intensive two-hour interviews, which served not only as an effective method for collecting data on the women served and the impact of the program, but also as an important part of the program intervention itself. These interviews became a time for participants to reflect on their lives, the paths they had taken, the barriers they had faced, and where they were now as participants in the program. The interview process also made them feel like valuable contributors to the evaluation process. One participant put it this way, "You really make us feel like we have something to offer … that you really care about our perspective and think we have something to say. It feels good." This seemed particularly important to program staff, given how disenfranchised and unempowered many of the women in the program felt.

By investing in the developmental aspects of evaluation, the evaluation team was able to transform the evaluation process from a negative experience to a positive and beneficial one, where staff, participants, volunteers, and other stakeholders were able to develop skills and capacities in the process of collecting important evaluative information.

Sources of Information About Evaluation

Brown, Lyn Mikel, and Carol Gilligan, "Listening for Self and Relational Voices: A Responsive/Resisting Reader's Guide," Monograph, Cambridge: Harvard Project on Women's Psychology and Girls' Development, 1990.

Brown, Lyn Mikel, and Carol Gilligan, *Meeting at the Crossroads: Women's Psychology and Girls' Development*, Cambridge: Harvard University Press, 1992.

Bryk, A.S. (ed.), *Stakeholder-Based Evaluation*, San Francisco: Jossey-Bass, 1983.

Campbell, D.T., and J.C. Stanley, *Experimental and Quasi-Experimental Designs for Research*, Chicago: Rand McNally, 1966.

Chen, Huey-tsyh, *Theory-Driven Evaluations*, California: Sage Publications, 1990.

Connell, P. James, Anne C. Kubisch, Lisbeth B. Schorr, and Carol H. Weiss, *New Approaches to Evaluating Communities Initiatives: Concepts, Methods, and Contexts*, Washington, DC: The Aspen Institute, 1995.

Covert, R.W., *Guidelines and Criteria for Constructing Questionnaires*, University of Virginia: Evaluation Research Center, 1977.

Cronbach, Lee J. and Associates, *Toward Reform of Program Evaluation,* San Francisco: Jossey-Bass, 1980.

Fetterman, David M., Shakeh J. Kaftarian, and Abraham Wandersman, (eds.) *Empowerment Evaluation: Knowledge and Tools for Self-Assessment & Accountability*, Thousand Oaks, CA: Sage Publications, 1996.

Feuerstein, Marie-Therese, *Partners in Evaluation*, London: Macmillan Publishers, 1986.

Fink, A., and J. Kosecoff, *How to Conduct Surveys: A Step-by-Step Guide*, Newbury Park, CA: Sage Publications, 1989.

Freeman, H.E., G.D. Sandfur, and P.H. Rossi, *Workbook for Evaluation: A Systematic Approach*, Newbury Park, CA: Sage Publications, 1989.

Frey, J., *Survey Research by Telephone*, 2nd ed. Newbury Park, CA: Sage Publications, 1983.

Furano, Kathryn, Linda Z. Jucovy, David P. Racine, and Thomas J. Smith, *The Essential Connection: Using Evaluation to Identify Programs Worth Replicating*, Replication and Program Strategies, Inc., Philadelphia, PA, 1995.

Gilligan, Carol, "Women's Place in Man's Life Cycle." In Sandra Harding (ed.) *Feminism and Methodology*, Bloomington: Indiana University Press, 1987.

Gilligan, Carol, Lyn Mikel Brown, and Annie Rogers, "Psyche Embedded: A Place for Body, Relationships and Culture in Personality Theory." In A. Rabin, R. Zucker, R. Emmons and S. Frank (eds.) *Studying Persons and Lives*, New York: Springer, 1990.

Gray, S.T. and Associates, *Evaluation With Power: A New Approach to Organizational Effectiveness, Empowerment, and Excellence,* San Francisco: Jossey-Bass Publishers, Inc., 1998.

Guba, E.C., and Y.S. Lincoln, *Effective Evaluation*, San Francisco: Jossey-Bass Publishers, Inc., 1981.

Henry, G.T., *Practical Sampling*, Newbury Park, CA: Sage Publications, 1990.

Herman. J.L., *Program Evaluation Kit*, 2nd ed., Newbury Park, CA: Sage Publications, 1987.

House, E.R., *Evaluating With Validity*, Beverly Hills, CA: Sage Publications, 1980.

Isaac, S., and W.B. Michael, *Handbook in Research and Evaluation*, San Diego: Edits Publishers, 1984.

Kidder, L.H., C.M. Judd, and E.R. Smith, *Research Methods in Social Relations*, 6th ed., New York: Holt, Rinehart & Winston, Inc., 1991.

Kosecoff, Jacqueline, and Arlene Fink, *Evaluation Basics: A Practitioner's Manual*, Newbury Park, CA: Sage Publications, 1982.

Krueger, R.A., *Focus Groups: A Practical Guide for Applied Research*, Newbury Park, CA: Sage Publications, 1988.

Lavrakas, P.J., *Telephone Survey Methods: Sampling, Selection, and Supervision*, Newbury Park, CA: Sage Publications, 1987.

Linney, J.A., and A. Wandersman, *Prevention Plus III: Assessing Alcohol and Other Drug Prevention Programs at the School and Community Level*, Rockville, MD: U.S. Dept. of Health and Human Services, 1991.

Littell, J.H., *Building Strong Foundations: Evaluation Strategies for Family Resource Programs*, Chicago, IL: Family Resource Coalition, 1986.

Love, A.J. (ed.), *Internal Evaluation: Building Organizations From Within,* Newbury Park, CA: Sage Publications, 1991.

Madaus, G., M. Scriven, and D.L. Stufflebeam, (eds.), *Evaluation Models*, Boston: Kluwer-Nijhoff Publishing, 1985.

Maxwell, J.A. and Y.S. Lincoln, *Methodology and Epistemology: A Dialogue,* Harvard Educational Review, 60(4), P.497-512, 1990.

McKnight, John L. and John Kretzmann, *Mapping Community Capacity*, Chicago: Center for Urban Affairs and Policy Research at Northwestern University, 1990.

Miles, M.B., and A.M. Huberman, *Qualitative Data Analysis: An Expanded Sourcebook*, 2nd ed., Thousand Oaks, CA: Sage Publications, 1994.

Morris, Lynn Lyons, and Carol Taylor Fitz-Gibbon, *How to Deal with Goals and Objectives*, Beverly Hills, CA: Sage Publications, 1978.

Morris, Lynn Lyons, and Carol Taylor Fitz-Gibbon, *How to Design a Program Evaluation*, Beverly Hills, CA: Sage Publications, 1978.

Morris, Lynn Lyons, and Carol Taylor Fitz-Gibbon, *How to Measure Program Implementation*, Beverly Hills, CA: Sage Publications, 1978.

Morris, Lynn Lyons, and Carol Taylor Fitz-Gibbon, *How to Present an Evaluation Report*, Beverly Hills, CA: Sage Publications, 1978.

National Center for Clinical Infant Programs, *Charting Change in Infants, Families, and Services: A Guide to Program Evaluation for Administrators and Practitioners*, Washington, DC: Author, 1987.

Nielson, Joyce M., "Introduction" in J. Nielson (Ed.) *Feminist Research Methods,* Boulder: Westview Press, 1990.

———, *Evaluating Service Programs for Infants, Toddlers, and Their Families: A Guide for Policy Makers and Funders*, Washington, DC: Author, 1987.

Patton, Michael Quinn, *Practical Evaluation*, Newbury Park, CA: Sage Publications, 1982.

———, *Qualitative Evaluation and Research Methods*, 2nd ed., Newbury Park, CA: Sage Publications, 1990.

———, *Utilization-Focused Evaluation: The New Century Text*, 3rd ed., Beverly Hills: Sage Publications, 1997.

Posavac and Carey, *Program Evaluation,* 5th ed., Prentice-Hall. 1997.

Reason, Peter, *Human Inquiry in Action: Developments in New Paradigm Research*, Newbury Park, CA: Sage Publications, 1993.

Reinharz, Shulamit, "Experimental Analysis: A Contribution to Feminist Research" in G. Bowles and R. Klien (Eds.), *Theories of Women's Studies,* Boston: Routledge and Kegan Paul, 1993.

Rossi, Peter H., and Howard E. Freeman, *Evaluation: A Systematic Approach*, Newbury Park, CA: Sage Publications, 1993.

Sanders, James R., *The Program Evaluation Standards: How to Assess Evaluations of Educational Programs*, 2nd ed., Thousand Oaks, CA: Sage Publications, 1994.

Saxe, L., and M. Fine, S*ocial Experiments: Methods for Design and Evaluation*, Beverly Hills: Sage Publications, 1981.

Schalock, Robert L., and Craige V.D. Thornton, *Program Evaluation: A Field Guide for Administrators*, New York: Plenum Press, 1988.

Schorr, Lisbeth B., and Anne C. Kubisch, "New Approaches to Evaluation: Helping Sister Mary Paul, Geoff Canada, and Otis Johnson While Convincing Pat Moynihan, Newt Gingrich, and the American Public," Presentation. Annie E. Casey Foundation Annual Research/Evaluation Conference: Using Research and Evaluation Information to Improve Programs and Policies, September 1995.

Scriven, Michael, *Evaluation Thesaurus*, 4th ed., Newbury Park, CA: Sage Publications, 1991.

Shadish, William R., Jr., Thomas D. Cook, and Laura C. Leviton, *Foundations of Program Evaluation: Theories of Practice*, Newbury Park, CA: Sage Publications, 1991.

Smith, M.F., *Evaluability Assessment: A Practical Approach*, Boston: Kluwer Academic Publishers, 1989.

Stainback, W. and S. Stainback, *Understanding and Conducting Qualitative Research*, Dubuque, IA: Kendall/Hunt Publishing Company, 1988.

Stockdill, Stacey Hueftle, *How to Evaluate Foundation Programs*, St. Paul, MN: The Saint Paul Foundation, Inc., 1993.

Travers, J.R., and R.J. Light, *Learning From Experience: Evaluating Early Childhood Demonstration Program*s, Washington, DC: National Academy Press, 1982.

United Way of America, *Current United Way Approaches to Measuring Program Outcomes and Community Change*, Alexandria, VA, 1995.

van der Eyken, W. *Introducing Evaluation*, The Hague: Bernard van Leer Foundation, 1992.

Weiss, Carol H., "Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families." In James Connell, et al. (ed.), *New Approaches to Evaluating Community Initiatives: Concepts, Methods and Contexts*, Washington D.C.: The Aspen Institute, 1995.

Weiss, H.B., and F.H. Jacobs, (eds.), *Evaluating Family Programs*, New York: Aldine de Gruyter, 1988.

Weiss, H.B., and R. Halpern, *The Challenges of Evaluating State Family Support and Education Initiatives: An Evaluation Framework*, Cambridge, MA: Harvard Family Research Project, 1989.

Wholey, Joseph S., Harry P. Hatry, and Kathryn E. Newcomer, (eds.), *Handbook of Practical Program Evaluation*, San Francisco, CA: Jossey-Bass Publishers, 1994.

Wilkin, David, Leslie Hallam, and Marie-Anne Doggett, *Measures of Need and Outcome for Primary Health Care*, Oxford: Oxford University Press, 1992.

Williams, Harold S., Arthur Y. Webb, and William J. Phillips, *Outcome Funding: A New Approach to Targeted Grantmaking*, 2nd ed., Rensselaerville, NY: The Rensselaerville Institute, 1993.

Worthen, Blaine R., and James R. Sanders, and Jody L. Fitzpatrick, *Program Evaluation: Alternative Approaches and Practical Guidelines*, 2nd ed., White Plains, NY: Longman Inc., 1997.

Yavorsky, D.K., *Discrepancy Evaluation: A Practitioner's Guide*, University of Virginia: Evaluation Research Center, 1984.

Yin, R.K., *Case Study Research: Design and Methods*, rev. ed., Newbury Park, CA: Sage Publications, 1989.

ACKNOWLEDGEMENTS

Director of Evaluation
Ricardo A. Millett (former)

Editorial Direction and Writing
Susan Curnan
Lisa LaCava
Dianna Sharpsteen
Mark Lelle
Michelle Reece

Production Assistance
Dianna Sharpsteen
Linda Tafolla

Design
Foursquare Communications

This *W.K. Kellogg Foundation Evaluation Handbook* is a reflection of the collective work of the Evaluation Unit. It has benefitted from the experiences of many evaluation practitioners and project directors. Much of the handbook is based on the *W.K. Kellogg Foundation Program Evaluation Manual*, compiled and written by James Sanders.

To give feedback on this publication, write to Teresa R. Behrens, W.K. Kellogg Foundation, One Michigan Avenue East, Battle Creek, MI, 49017-4058, USA or send e-mail to *evaluation@wkkf.org*.

Hard copies of this *Evaluation Handbook* are available free of charge by phone at 800-819-9997 or written inquiry to:

W.K. Kellogg Foundation
P.O. Box 550
Battle Creek, MI 49016-0550

Some limitations on quantities apply.

The *W.K. Kellogg Evaluation Handbook*, and other evaluation materials, can also be viewed on-line at *www.wkkf.org.*